

PERCEPTUAL SENSE UNITS IN THE PROCESS OF LISTENING COMPREHENSION

Morio Kohno

Kobe City University of Foreign Studies
Kobe, Japan.

ABSTRACT

Research was added to the previous studies which made clear that the perception of sound sequence consists of two-fold processings--holistic and analytic, and that the former is applied to the fast tempo sequences whose intersound intervals are within 300ms, and the latter to the slow ones whose intervals are over 400ms. They are neuropsychologically different from each other (See another paper of the author in this conference named "Two Processing Mechanisms in Rhythm Perception.") Experiments I and II proved that in the process of listening comprehension, meaning units which are made up of 1 to 7 ± 2 syllables closely combined each other with the intervals of less than 300ms are perceived holistically as definite units. If holisticality is deprived semantically and/or physically by pausing, the listenability is extraordinarily decreased. There is some evidence that this 'perceptual sense unit' was imbedded in human-beings' deep cognitive system.

The present paper is going to try to identify 'perceptual sense unit' (P-unit, henceforth), perceived holistically and stored in echoic memory in an unprocessed form in the process of listening comprehension.

1. PREVIOUS STUDIES

Kohno and Kashiwagi [3] made clear the processing mechanism in rhythm perception using as subjects a normal right-handed woman, seven children with age variety from one

year and four months to nine years old and a patient with infarction involving the forebrain commissural fibers. The summary of the results got by the study is as follows.

1) The left hand of the patient and children under four years of age can not synchronize their tapping with the slow rhythms of 500 and 1000ms inter-beat intervals (IBI), but they can follow the rapid stimuli of 250ms IBI. The right hand of the patient and children older than four years old, however, can fit their tapping both to the rapid and slow rhythms as well as the normal adult. 2) Negative autocorrelations were detected among adjacent IBIs in slow response beats by the normal adult, the children older than four and the right hand of the patient, but never found in the responses to the slow stimuli by the children younger than four and the left hand of the patient. 3) Negative autocorrelations were never detected in the rapid response movements (250ms) of all the subjects. 4) The above-mentioned facts suggest that the slow repetitive sound sequences are normally perceived by ongoing and analytic way of processing, but rapid ones by at-a-time and holistic way. Evidence was found that the children younger than four years old and the left hand of the patient always use only the holistic approach not only for the rapid rhythm but also for the slow rhythm and that it is the very reason why they cannot synchronize the slow tempos. These two kinds of processing, therefore, are neuropsychologically different from each other. 5) The above-mentioned facts suggest that the slow repetitive sound sequences are normally perceived by ongoing and analytic way of processing, but rapid ones by at-a-time and holistic way. Evidence was found that the children younger than four years old and the left hand of the patient always use only the holistic approach not only for the rapid rhythm but also for the slow rhythm and that it is the very reason why they cannot synchronize the slow tempos. These two kinds of processing, therefore, are neuropsychologically different from each other. 6) Other experiments using

nonsense words on timing condition of syllable sequences and echoic memory were held with the following results. a) The sequences of closely connected nonsense syllables, each of whose inter-voice-onset intervals (IVI, henceforth) are less than 300ms, bring forth longer retention than the syllable sequences whose IVIs are longer than 400ms, when they are recalled after some lapse of time for doing two digit number multiplication. As suggested in 5), the durational condition less than 300ms may be processed holistically, and that more than 400ms, analytically. As holistic processing is qualitatively different from analytic one, the former is never disturbed by the latter, and this may be the reason why the words whose syllables are closely connected have longer retention than the ones which consist of loosely connected syllables after doing some cognitive work. b) There is some evidence that durational condition between 300 and 400ms IVIs is border area mixed with both holistic and analytic processings, different by individuals.

2. IDENTIFICATION OF PERCEPTUAL SENSE UNIT

2.1. Experiment I

[Subjects] Students of a high school in Japan, 108 in number, were divided into four homogeneous groups in their English ability based on their academic records.

[Materials] Two original stories, one in English (95 words), the other, in Japanese (133 words), were recorded by an American instructor and a Japanese one respectively. Then pauses were mechanically placed by the use of Pause Controller, SONY LLC 5000, at every end of word (the set pause length = 1 second), at every end of phrase, of clause and of sentence (the set pause length = 2 seconds). Phrase here means a meaning unit which consists of one content and no or some function words. A no-pause version was also prepared. An important thing here is that the number of syllables which made up a phrase was 1 to 7 in English version and 2 to 8 in Japanese. The unit

of clause, however, consisted of 2 to 13 syllables in English, and 7 to 25 (8 to 26 morae) in Japanese, and the sentence unit, 12 to 22 in English and 12 to 48 (15 to 53 morae) in Japanese. According to Miller [4], maximum number of elements which can be perceived in a flash is 7 ± 2 . Only the unit of phrase out of the above units meets Miller's condition--in the case of English clause, for example, half of them consisted of more than 7 syllables. Words, of course, consist of less than 7 syllables, but the separation by this unit destroys the unit of meaning, because no function word has any independent meaning.

[Method] Each subject group was requested to listen to one version in Japanese and another in English, and were asked to write the content of the stories as precisely as possible in Japanese, immediately after they had finished hearing them.

Table 1 - A (Japanese) full marks = 29

| pause | n | \bar{x} | S.D. |
|-----------------|----|-----------|------|
| no pause(A) | 27 | 4.2 | 1.78 |
| every clause(B) | 26 | 9.7 | 5.27 |
| every phrase(C) | 24 | 15.6 | 5.65 |
| every word(D) | 31 | 11.8 | 5.82 |

B > D ($t=1.39$, N.S.); C > D ($t=2.39$, $P<0.05$)
C > B ($t=3.79$, $P<0.001$); C > A ($t=9.75$, $P<0.001$)
B > A ($t=2.55$, $P<0.02$); D > A ($t=5.41$, $P<0.001$)
(Table 1 - B (English) omitted.)

[Results] Pauses which were placed at every end of clauses, and of phrases increased the scores in this order in both English and Japanese materials--the no-pause version produced the worst ones. As found in [2], the pause-at-every-phrase version brought about the highest scores. This 'the more pauses, the higher score' principle, however, did not go on in the pause-at-every-word versions, which remarkably reduced the scores ($p<0.05-0.001$). This fact also coincides with the result of [3]. If pauses would play a crucial role for listening comprehension by giving chances to analyze and synthesize the stimuli and by giving clues to separate the sound stream into proper units as pointed out in Pimsleur (1971), we may rightly say that these results suggest

the unit of phrase, more precisely, grammatical meaning unit which consists of 7 ± 2 syllables, might render a most suitable chunk for listeners' cognitive processing of connected speech.

We should now notice that the P-unit, the unit of phrase, for example, generally consists of several syllables which are combined one to another with IVIs about 100 to 200ms in the case of Japanese or 100 to 250ms in English, all of which are so rapid as to be processed holistically. The last IVIs of each P-unit in English, however, are somewhat longer than the preceding ones. Here is an example of realities of IVIs together with syllable lengths, that is, computational observations (ILS, Micro PDP 11/73) on the syllable durations and IVIs between syllables in the stories read by native speakers.

a) Duration of syllables (syl. dur.) and IVIs in spoken Japanese sentences.

Tsu ki ga / no bo ri/
 syl.dur. 118 104 96 108 111 115
 IVI * 146 143 121 116

ha ji me ma shi ta./
 syl.dur. 139 139 163 165 192 158
 IVI 169 111 143 (117) (185)

b) Duration of syllables (syl. dur.) and IVIs in spoken English sentences

Scott/ came out of the
 syl.dur. 331 225 242 115 196
 IVI 386 225 221 165 148

house/ and locked / the
 syl.dur. 311 122 345 (183) 168
 IVI 232 290

door / behind him.
 syl.dur. 329 162 308 345
 IVI 170 287 198 323

= boundary of p-unit)

In Japanese, all IVIs, as well as syllable durations, are all within 300 ms. This means that the syllable sequences so closely connected have to be processed holistically. In English, however, the states of IVIs and syllable durations vary very much, and generally they are longer than Japanese syllables. We can notice, however, that in English, a long IVI is put at each

P-unit, especially at the end of it, that is, at each semantic unit which usually consists of 7 or less syllables. The fore parts of the unit are composed by a syllable succession with short IVIs which may be processed holistically. To put the long syllables at the ends of units is very effective to show the terminations of the units. This device seems to help listeners out of difficulty of holistic processing caused by the variety of syllable durations.

We can now more precisely say that a perceptual sense unit, is a semantic unit which is composed of one to several syllables which are so closely combined one to another in less than 300ms IVIs so that the unit can be processed holistically.

If the unity and, therefore, holisticness are lost by the long IVIs of more than 500ms, listenability of the utterance will be remarkably decreased. The analytic processing, on the other hand, may concern the processing of two and more perceptual sense units one by one to get the whole meaning of utterances. It might, as a matter of course, take a longer time.

2.3. Experiment III

In order to verify the above hypothesis, the following experiment was carried out.

[Subjects and Method] An essay in Japanese (Material A) was read by a Japanese female instructor (age: twenties). It was then mechanically separated at every end of perceptual sense units by pauses whose durations were 3, 4 and 5 seconds. Another Japanese essay (Material B) was also read by the same instructor, but IVIs among syllables were spread to each of 200, 250, and 500ms. Subjects were Japanese high school students, 122 in number. Other procedures were the same as in Experiment I.

[Results and Discussion] The results are shown in Tables 2-A and -B. When the IVI among syllables was 200 or 250ms which will be processed holistically, the scores remained almost the same, but when they were

lengthened to 500ms, which may be processed analytically, the scores significantly decreased (Table 2-A).

Table 6 - A

| | Intervals among syllables | | |
|-----------|---------------------------|-------|-------|
| | 200ms | 250ms | 500ms |
| n. | 27 | 34 | 26 |
| \bar{x} | 14.52 | 13.03 | 11.27 |
| S. D. | 2.49 | 3.91 | 4.52 |

200ms \approx 250ms N. S.; 200ms > 500ms $P < 0.01$

Table 6 - B

| | Intervals among phrases | | | |
|-----------|-------------------------|--------|--------|--------|
| | natural | 3000ms | 4000ms | 5000ms |
| n. | 25 | 26 | 34 | 27 |
| \bar{x} | 15.88 | 16.07 | 16.38 | 16.11 |
| S. D. | 2.09 | 2.00 | 1.67 | 1.95 |

NS among factors⁵

In connection with this, [4] pointed out that extremely slow English pronunciation in which each syllable is drawn more than 500ms ($\bar{x}=514$ actually) also decreases listenability very much ($p < 0.05-0.01$), even though pronunciation of segments was clear and precise. Table B, on the other hand, shows that long intervals among the perceptual sense units never bring about negative effects, perhaps because the work to be done here would be an analytic task to look for grammatical and semantic relations among several units, which by nature requires lots of time.

3. DISCUSSION

We can now exactly explain the results of Experiment I: the no-pause version brought about the worst scores, because all the chances to perform the work of analytic processing were deprived. The pause-at-every-word version also produced very bad scores, because the unity of meaning is lost. The low listenability of drawled pronunciation [3] can be explained in the same way. Perceptual sense unit is a unit in which 'a unit of meaning' is closely attached to a closely connected syllable succession. Listeners, there-

fore, when they hear the unit, can reflectively recall the meaning to their mind. In the pause-at-every-word versions, however, function words which are put separately by pauses whose durations are beyond holistically perceived time intervals, prevent listeners from reflexive, at-a-time recalling of a unit of meaning. The relatively low scores produced by the pause-at-every-sentence version and by the pause-at-every-clause version can be explained by the fact that the number of syllables composing those units go beyond the suitable number that human beings can holistically process at a time, that is, 7 ± 2 .

4. CONCLUSION

We may conclude by saying that the processing of listening comprehension is a mix of both holistic and analytic works. The existence of the unit which should be holistically dealt with at a time has long been overlooked, but it is crucial element to make clear the processing of listening comprehension which listeners can do very efficiently and rapidly. We should notice that perceptual units whose syllables are closely connected may be preserved in an unprocessed form longer than the separately lined syllables after the execution of some cognitive works [3]. On account of this nature, listeners can do the works of semantic and sometimes grammatical analysis over several units, if necessary, referring back to some precedent unit which is still retained even after having processed some units.

REFERENCES

- [1] KOHNO, M. (1981), "Effect of pause on listening comprehension." T. Konishied. *Studies in Grammar and Language*, 392-405, Kenkyusha, Tokyo.
- [2] KOHNO, M. Kashiwagi, A. and Kashiwagi, T. (1990), "Similarity of rhythms produced by young children and the patient with infarction involving the corpus callosum." *AILA* 1990.
- [3] MILLER, G. (1956), "The magical number seven, plus or minus two." *Psychological Review*, 63-2, 81-97.