DOI: 10.61782/fa.2023.0070

forum acusticum 2023

EFFECTS OF SPEAKING RATE AND INDIVIDUAL DIFFERENCES ON PERCEPTUAL TRANSITIONS

Seiya Funatsu*

Science Information Center, Prefectural University of Hiroshima, Japan

ABSTRACT

The perceptual transition occurs when a person hears a repeated single word without pause, illusory changes of the physically unchanging word are induced. Our research laboratory have been studying the conditions under which perceptual transitions occur. This study investigated the effect of speaking rate on the occurrence of perceptual transitions. This study also examined the effects of individual differences in subjects on the occurrence of perceptual transitions. The time until the perceptual transition occurred by varying the speaking rate. The speaking rates were fast rate (average: 7.85 syllables/second) and slow rate (average: 5.67 syllables/second). The stimulus consisted of Japanese words with two, three and four syllables. The subjects were nine native Japanese speakers. They clicked a button on a PC screen when they perceived a change in the sound. As a result, there was no significant difference in the likelihood of perceptual transitions due to changes in speaking rate. Since speaking rate did not affect how perceptual transitions occurred, it was inferred that perceptual transitions were not caused by habituation, neural fatigue, or saturation/overflow of short-term memory. Furthermore, individual differences varied depending on the combination of speaking rate and stimuli, and in some cases the individual differences were relatively large, while in other cases they were relatively small. Within the scope of this experiment, it was suggested that the likelihood of occurrence of perceptual transitions is robust against changes in speaking rate.

*Corresponding author: funatsu@pu-hiroshima.ac.jp.

Keywords: *verbal transformation effect, perceptual transition, speaking rate, 2-4 syllable words, Japanese words.*

1. INTRODUCTION

When one word is presented repeatedly without pause, the perceptual transition occurs, and subjective changes of percepts arise. This phenomenon is called the verbal transformation effect (VTE).

Since VTE was first reported by Richard M. Warren et. al. in 1958 [1], their investigation has focused on English [2], [3]. For example, “tress” may be transformed into a variety of verbal forms, such as “dress,” “stress,” “drest,” or even “Esther” [2], [3]. Studies in Japanese have explored brain responses, and one such study by Kondo and Kashino (2007) clarified brain activity during perceptual transitions using fMRI [4].

In previous studies by Funatsu & Fujimoto, they have found that perceptual transitions occur not only without pause, but when inter stimulus interval (ISI) is about 400 ms [5]. Moreover, they found that the longer the ISI, the longer it takes for the perceptual transition to occur (the number of perceived stimuli increases before the perceptual transition occurs) [6], [7]. However, it is not clear at what speaking rate the perceptual transition occurs.

Perceptual transitions are said to be caused by habituation or neural fatigue due to repeatedly presented stimuli [8], [9]. If perceptual transitions are caused by habituation, neural fatigue, or saturation/overflow of short-term memory, they should be affected by speaking rate. In other words, when

the speaking rate is fast, habituation, neural fatigue, or saturation/overflow of short-term memory occur in a short time, and when the speaking rate is slow, the time until habituation, neural fatigue, or saturation/overflow of short-term memory occur is long. In this study, we investigate the possibility that perceptual transitions are caused by habituation, neural fatigue, or short-term memory saturation/overflow by changing the speaking rate.

2. METHODS

2.1 Stimuli

The stimuli used in this experiment were speech sounds. These stimuli were two-syllable, three-syllable, and four-syllable words. Three-syllable words are two-syllable words plus one syllable, and four-syllable words are three-syllable words plus one syllable. The stimuli are the following six Japanese words. Group A words are words based on CV syllables ($C_1V_1C_2V_2$, $C_1V_1C_2V_2C_3V_2$, $C_1V_1C_2V_2C_3V_2C_4V_2$). Words in Group B are vowel-based words (V_1V_2, $V_1V_2V_3$, $V_1V_2V_3C_1V_4$).

Group A
/tera/ (temple)
/teraja/ (a private elementary school in the Edo period)
/terajama/ (family name)

Group B
/ao/ (blue (noun))
/aoi/ (blue (adjective))
/aoiro/ (blue color)

The above words were uttered by a female standard Japanese speaker from the south Kanto region at two different speaking rates, fast and slow. The average of fast speaking rate was 7.85 syllables/second and that of slow rate was 5.67 syllables/second. The following table shows the speaking rate of individual words.

Table 1. Speaking rates of words. (syllables/second)

stimuli	fast rate	slow rate
/tera/	8.00	5.63
/teraja/	7.65	5.46
/terajama/	8.00	5.97
/ao/	7.14	5.13
/aoi/	7.69	6.20
/aoiro/	8.64	5.63
Mean	7.85	5.67

2.2 Subjects and procedure

The subjects were 12 native Japanese speakers aged 20-22 (12 females). However, since the sound stream segregation occurred in three subjects, they were excluded from the analysis. Therefore, the actual number of subjects analyzed was 9.

The stimulus was presented on a PC, and the subjects listened to the stimulus sequence through headphones (Sennheiser HDA200). This perception test was conducted individually in a quiet room. Each stimulus sequence, fast, slow, was presented for 90 s. They were instructed to click a button on a PC screen when they perceived a change in the sound. Start time of sequences (t_0) and click time (t_c) were recorded on a computer, and the perceptual transition time (T) was calculated by subtracting t_0 from t_c.

$$T = t_c - t_0 \qquad (1)$$

However, the longer the ISI, the longer the time until the perceptual transition occurs. Therefore, the number of stimuli heard by the subject until the perceptual transition occurred was used instead of time.

$$Number\ of\ stimuli = T / (stimulus\ length) \qquad (2)$$

These stimuli were presented in order of fast rate and slow rate. And they were presented in the order shown in Table 1. Previous studies by Funatsu & Fujimoto have shown that the results do not depend on the order of stimulus presentation [7], [10]. After perception tests were completed, we also asked each subject what they perceived the sound to be.

3. RESULTS

Perceptual transitions did not occur in some stimulus sequences for several subjects. In these cases, the perceived number of stimuli was calculated by dividing the length of the stimulus sequence, i.e. 90 s, by the length of each stimulus. These cases are shown in italics in Tables 3 and 4.

Table 2 shows the results of the experiment. This table shows the number of stimuli each subject heard before the perceptual transition occurred. This is the average of all stimuli at fast and slow speaking rates. The results varied widely from subject to subject, with subject A having almost no difference in speaking rate, subject C having more stimuli heard at fast rate than at slow rate, and conversely, subject F having more stimuli heard at slow rate than at fast rate.

Table 2. Number of stimuli perceived at fast and slow speaking rates.

subject	fast	slow
A	36.3	36.5
B	46.2	40.0
C	118.9	65.8
D	111.9	106.5
E	31.3	56.6
F	57.7	108.7
G	50.6	54.3
H	124.5	87.7
I	91.3	85.8

Table 3 shows the results for Stimulus Group A. This table shows the number of stimuli heard before the perceptual transition occurred for each subject. A two-way ANOVA was conducted with speaking rate and stimulus as factors.

The main effect of speaking rate was not significant ($F(1,8)=1.479$, $p=0.259$). The main effect of stimulus was significant at the 5% level with $F(2,16)=5.947$, $p=0.0117$. The interaction was significant at the 1% level with $F(2,16)=8.604$, $p=0.0029$. The standard deviations and coefficients of variance were calculated to examine differences between subjects. In Group A, the standard deviations were large and the coefficients of variance exceeded 0.7 for the /tera/ fast speaking rate and /teraja/ slow speaking rate. Individual differences were considered to be relatively large for these stimuli and speaking rates. On the other hand, the coefficient of variance for /terajama/ was relatively small, less than 0.5 for both fast and slow speaking rates. In this case, individual differences were considered to be small.

Table 4 shows the results for Stimulus Group B. A two-way ANOVA was conducted with speaking rate and stimulus as factors. The main effect of speaking rate was not significant ($F(1,8)=2.562$, $p=0.148$). The main effect of stimulus was significant at the 1% level with $F(2,16)=8.876$, $p=0.0026$. The interaction was significant at the 5% level with $F(2,16)=4.24$, $p=0.033$. The coefficient of variance in Group B exceeded 0.7 for the slow speaking rate of /aoi/, but the coefficient of variance was less than 0.4 for the fast speaking rate of /aoi/ and the slow speaking rate of /aoiro/, suggesting that individual differences were relatively smaller for the Group B stimuli than for the Group A stimuli.

Table 3. Number of stimuli perceived in Stimulus Group A.
S.D.: standard deviation, C.V.: coefficient of variance

	/tera/		/teraja/		/terajama/	
subject	fast	slow	fast	slow	fast	slow
A	58.4	36.1	57.4	43.4	29.8	53.4
B	104.0	24.5	67.9	29.0	28.4	76.0
C	*360.0*	104.8	103.6	20.6	60.4	23.3
D	*360.0*	253.5	54.6	35.9	55.2	*134.3*
E	48.0	65.9	28.1	33.0	19.0	26.7
F	82.0	82.8	73.5	*163.9*	41.4	60.3
G	87.6	70.7	37.0	74.0	39.6	65.2
H	*360.0*	121.7	168.9	*163.9*	40.8	106.4
I	183.6	104.5	53.1	119.9	86.0	70.4
Mean	182.6	96.1	71.5	75.9	44.5	68.5
S.D.	130.5	63.3	40.0	55.0	19.0	33.2
C.V.	0.715	0.659	0.560	0.724	0.428	0.485

Table 4. Number of stimuli perceived in Stimulus Group B.
S.D.: standard deviation, C.V.: coefficient of variance

	/ao/		/aoi/		/aoiro/	
subject	fast	slow	fast	slow	fast	slow
A	17.1	21.8	14.9	28.1	40.4	36.1
B	30.7	17.7	29.5	36.4	16.8	56.4
C	96.1	55.6	69.5	63.6	23.8	*126.6*
D	32.9	47.4	73.8	41.1	94.8	*126.6*
E	21.4	22.6	48.2	64.9	22.9	*126.6*
F	40.7	32.6	55.4	*186.0*	53.3	*126.6*
G	15.4	14.4	50.3	37.2	74.1	64.3
H	49.3	20.8	33.3	48.3	94.6	65.3
I	102.1	37.2	42.3	56.2	80.8	*126.6*
Mean	45.1	30.0	46.4	62.4	55.7	95.0
S.D.	30.7	13.4	17.8	45.3	29.5	36.2
C.V.	0.681	0.448	0.384	0.725	0.530	0.381

Summarizing the above results, no main effect on speaking rate (fast vs. slow) was observed in both Group A and Group B. The main effect on stimulus was significant at the 5% level in Group A. This is probably due to the fact that the perceptual transitions did not occur at the fast speaking rate of /tera/ for 4 subjects. Group B was significant at the 1% level. This is also considered to be due to the fact that the perceptual transitions did not occur at the slow speaking rate of /aoiro/ for 5 subjects.

4. DISCCUSSION

There were no significant differences with respect to speaking rate for both Stimulus Group A and Stimulus Group B. That is, there was no difference in how the

perceptual transition occurred whether the speaking rate was fast or slow.

It is possible that repeated stimulation may cause habituation or neural fatigue, resulting in perceptual transition. However, in that case, it would be expected that the lower the speaking rate, the longer the time before the perceptual transition occurs. However, this possibility is unlikely, since no perceptual transitions were delayed at lower speaking rates.

If perceptual transitions occur in relation to short-term memory, i.e. if they are caused by the overflow/saturation of short-term memory, then it would also be expected, as with the case of habituation or neural fatigue, that the lower the speaking rate, the longer the time until the perceptual transition occurs. However, since this was not the case in the present experiment, it is unlikely that short-term memory is involved. These results are consistent with those of Kaminska et al. [8]. These results may suggest that perceptual transitions are not caused by habituation, neural fatigue, or short-term memory saturation/overflow in perception of Japanese words.

However, the difference between fast and slow speaking rates in this experiment was small, about 2 syllables/second, so it is possible that no perceptual transition delay was observed. Different results may be obtained when the experiment is conducted with stimuli with a larger difference in speaking rate.

5. CONCLUDING REMARKS

There was no significant difference in the occurrence of perceptual transitions due to changes in speaking rate. Within the scope of this experiment, it was suggested that the occurrence of perceptual transitions is robust against changes in speaking rate. Further studies using synthesized speech and strict control of speaking rate would be needed.

6. ACKNOWLEDGMENTS

My special thanks are due to Prof. Miyoko Ishige (Kyorin University) for her cooperation. And I also thank Prof. Xiao YeGui, Prof. Sumi Sugiyama and Prof. Hisako Orimoto (Prefectural University of Hiroshima) for their assistance. Finally, I am very grateful to Dr. Masako Fujimoto (former researcher at Waseda University) for her help throughout the research.

7. REFERENCES

[1] R.M. Warren, R.L. Gregory, "An auditory analogue of the visual reversible figure," *Am. J. Psychol*. 71, pp. 612–613. 1958.

[2] R.M. Warren, "Illusory changes of distinct speech upon repetition-The verbal transformation effect," *Br. J. Psychol*. 52, pp. 249–258. 1961.

[3] R.M. Warren, "Verbal transformation effect and auditory perceptual mechanisms," *Psychol. Bull*. 70, pp. 261–270. 1968.

[4] HM. Kondo, M. Kashino, "Neural mechanisms of auditory awareness underlying verbal transformations," *Neuroimage*. 36, pp. 123-130. 2007.

[5] S. Funatsu, M. Fujimoto, "Working of auditory short-term memory during continuous sound perception," *The 2020 spring meeting of the acoustical society of Japan*. (in Japanese) 2020.

[6] S. Funatsu, M. Fujimoto, "Fundamental research on the verbal transformation effect in Japanese," *Proc. 23rd International Congress on Acoustics, integrating 4th EAA Euroregio*. 2019.

[7] S. Funatsu, "Effect of inter-stimulus interval on verbal transformation effect," *The 64th Annual Meeting of the Japan Society of Logopedics and Phoniatrics*. (in Japanese) 2019.

[8] Z. Kaminska, M. Pool, P. Mayer, "Verbal transformation: Habituation or spreading activation?" *Brain and Language*. 71, pp. 285-298. 2000.

[9] K. A. Snyder, R. S. Calef, M. C. Choban, "Effects of word repetition and presentation rate on the frequency of verbal transformations: Support for habituation," *Bulletin of the Psychonomic Society*. 31, pp. 91-93. 1993.

[10] S. Funatsu, M. Fujimoto, "Basic research of verbal transformation effect," *The 2019 spring meeting of the acoustical society of Japan.* (in Japanese) 2019.