THE VERBAL TRANSFORMATION EFFECT IN JAPANESE

Seiya Funatsu*, Masako Fujimoto**

*Prefectural University of Hiroshima, **Waseda University
funatsu@pu-hiroshima.ac.jp

ABSTRACT

The verbal transformation effect is a phenomenon such that when a person hears a repeated single word without pause, illusory changes of the physically unchanging word are induced. While this effect has been studied in English and French, few studies have been undertaken about this effect in Japanese, especially from the perspective of behavioral science. In this study, we examined whether the verbal transformation effect occurs with or without pauses, and whether nonsense words also induce the verbal transformation effect. Our results, unlike those of previous studies, showed that the verbal transformation effect increased with a 0.15 sec pause. Moreover, the perceptual transition time was significantly longer with the 0.15 sec pause than without it. We found that the verbal transformation effect occurred with nonsense words, and also, the perceptual transition time varied depending on the word. The mechanisms of the verbal transformation effect are discussed.

Keywords: Inter stimulus interval, perceptual transition, Japanese, delay, illusory words.

1. INTRODUCTION

Listening to repetitions of a single word without a pause induces illusory changes of the physically unchanging word. For instance, “tress” may be transformed into a variety of verbal forms, such as “dress”, “stress”, “drest”, or even “Esther” [3, 6]. This phenomenon is known as the verbal transformation effect (VTE) [3]. Warren examined the influence of F0, loudness, and noise on the VTE in English [3].

A few studies on the VTE have been conducted from the perspective of brain science in Japanese. Using fMRI, Kondo & Kashino in their study revealed that the left inferior frontal cortex, anterior cingulate cortex, and the left prefrontal cortex were activated during the perceptual transition [2].

Japanese words are constructed by open syllables in general. Therefore, Japanese syllable orders can be for the most part exchanged and phonemes can be generally replaced by other phonemes. Few studies to date, however, have examined the VTE from the perspective of behavioral science in Japanese. This is a preliminary study to analyze the mechanisms of VTE from the perspective of behavioral science. Specifically, we tested the effect of satiation and temporal masking.

The perception of speech is satiated due to repeated activation. We examined whether perceptual transitions occur with and without pauses. In addition, because the state of satiation would change due to pauses, we also measured how long it takes for perceptual transitions to occur.

In the stream of repeated words, word-initial syllables would mask word-final syllables. The intensity of a word-initial syllable is stronger than that of a word-final syllable, thus backward masking would occur in the perception of repeated words. Moreover, we examined whether nonsense words induce perceptual transition.

2. METHODS

2.1. Stimuli

The stimuli were /banana/, /bamana/ and /nanaba/, uttered by a male native Japanese speaker. /banana/ is a real Japanese word for which VTE was found to occur (Kondo & Kashino [2]), while /bamana/ and /nanaba/ are nonsense words. In order to test the masking effect, two pronunciations of /banana/, normal and emphasized, were carried out. In the normal utterance (normal /banana/), the initial /ba/ in /banana/ had the strongest intensity and the final /na/ had the weakest intensity. In the emphasized utterance of /banana/ (emphasized /banana/), the final /na/ was emphasized and had almost the same intensity as the initial /ba/ (Fig. 1).

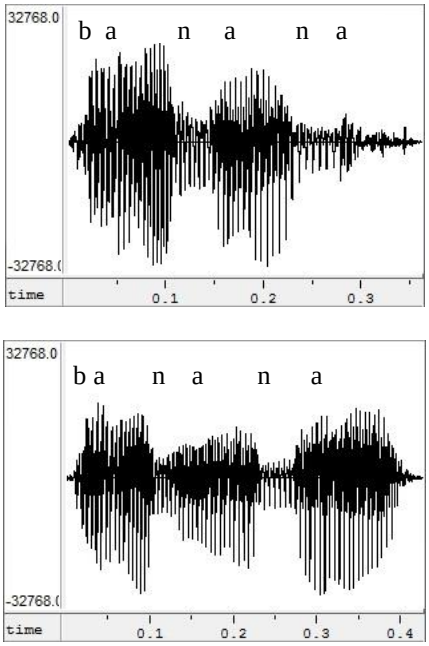
Two sequences of normal /banana/ were carried out. In one, normal /banana/ was repeated for 90 sec with no gaps (the inter stimulus interval (ISI) was 0 sec), and in the other, the ISI was 0.15 sec. The remaining stimuli, (emphasized /banana/, /bamana/, /nanaba/), were also repeated for 90 sec with no gaps.

2.2. Subjects and procedure

Subjects were 12 native Japanese university students aged 21-23 (male: 4, female: 8). They were instructed to listen to the stimulus sequences and then click a button on a computer display when they

perceived a change in the sound. They heard five sequences through headphones. In order to avoid the influence of the presentation order (ISI 0 sec, ISI 0.15 sec), subjects were divided two groups. Each group consisted of six subjects.

Figure 1: Wave forms of normal /banana/ (upper) and emphasized /banana/ (lower)



- The ascending order group (ASC) heard the stimulus sequences in the following order: /banana/ with a 0 sec ISI, /banana/ with a 0.15 sec ISI, emphasized /banana/, /bamana/, and /nanaba/.
- The descending order group (DSC) heard stimulus sequences in the following order: /banana/ with a 0.15 sec ISI, /banana/ with a 0 sec ISI, emphasized /banana/, /bamana/, and /nanaba/.

Start time of sequences (t_0) and click time (t_c) were recorded on a computer, and the perceptual transition time (PTT) was calculated by subtracting t_0 from t_c. Obtained PPTs from individual subjects were averaged for each stimulus.

After perception tests were completed, we also asked each subject what they perceived the sound to be.

3. RESULTS

3.1. ISI effects

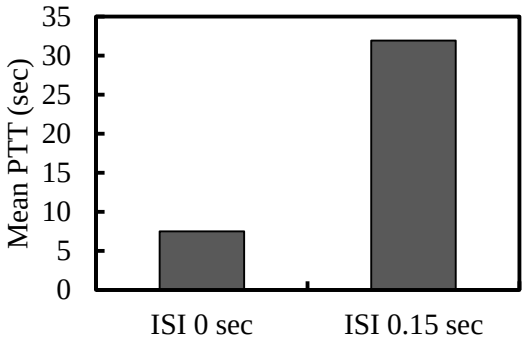
In the case of the 0.15 sec ISI, perceptual transition did not occur for one subject in the DSC. Table 1 shows the results for the ASC and DSC. In the case when ISI was 0 sec, the mean PTT for the ASC was 7.18 sec, and the mean PTT of the DSC was 7.80 sec. In the case when the ISI was 0.15 sec, the mean PTT of the ASC was 32.38 sec, and the mean PTT of the DSC was 31.38 sec. ANOVA with two factors, ISI and Order, showed a significant main effect of ISI ($F(1,19) = 13.14$ ($p = 0.002$)), but there were no main effects of Order ($F(1,19) = 0.001$ ($p = 0.977$)) and interaction ($F(1,19) = 0.014$ ($p = 0.905$)). From these results, it is clear that these two groups showed no difference in the mean PTT.

Figure 2 shows the ISI effects for all subjects. /banana/ with a 0.15 sec ISI needed a longer time for the perceptual transition to occur than that with 0 sec ISI.

Table 1: Mean PTT of ASC and DSC (in sec).

	ISI 0	ISI 0.15
ASC	7.18	32.38
DSC	7.80	31.38

Figure 2: The effect of ISI in normal /banana/ for all subjects.



3.2. Real words vs. nonsense words

Figure 3 shows the results for real words vs. nonsense words. Perceptual transition occurred even with nonsense words. As Fig. 3 shows, the mean PTTs for emphasized /banana/, /bamana/, and /nanaba/ were longer than the mean PTT for normal /banana/. ANOVA showed a significant difference among these words ($F(3,42) = 5.259$ ($p = 0.004$)). A *post hoc* test was performed and revealed a significant difference between normal /banana/ and /nanaba/ ($p = 0.0003$). /nanaba/ is a nonsense word in which the syllables of /banana/ are transposed. No other significant differences were found.

3.3. Illusory words in perceptual transition

Table 2 shows the illusory words that were perceived by each subject when perceptual transition occurred. When the ISI was 0 sec, six subjects perceived “banan”, and two subjects “banano”.

When the ISI was 0.15 sec, two subjects perceived "banan" and two subjects "wanana". Moreover, auditory stream segregation occurred for one subject. She perceived /banana/ (0.15 sec ISI) as two separate perceptual streams (/nba/ and /na/). When the ISI was 0.15 sec, the words that subjects perceived differed from those perceived when the ISI was 0 sec.

Figure 3: Mean PTT for each word. normal: normal /banana/; emphasized: emphasized /banana/.

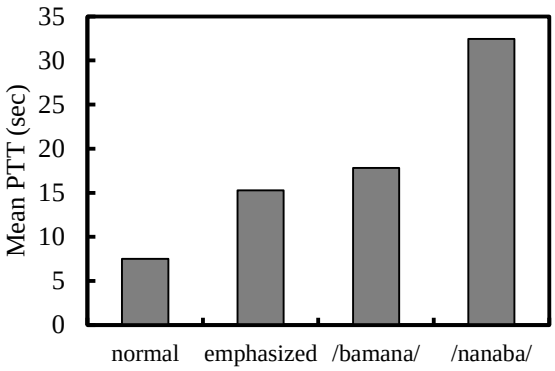


Table 2: Perceived illusory words for each subject.

Subject	Order	ISI 0	ISI 0.15
A	ASC	banan	banaka
		tanan	kanano
B	DSC	banan	wanana
		nanto	
C	ASC	banan	wanana
		nanpa	tarano
			tanano
D	DSC	tonen	ASS
E	ASC	banan	warana
			baran
F	DSC	banano	banan
		barano	
		bararo	
		banono	
G	ASC	banan	warano
H	DSC	banano	banan
			nanpa
I	ASC	banan	narano
		bananu	
J	DSC	kadan	*
K	ASC	ganani	yaneno
			porando
L	DSC	banan	**
		tanan	

ASS: Auditory stream segregation occurred.
*: The sound that subject D heard was indistinct.
**: Perceptual transition did not occur in ISI 0.15 sec.

3. DISCUSSION

Perceptual transition occurred when the ISI was 0.15 sec. As the ISI lengthened, so did the PTT. Stimulus length of normal /banana/ was 0.365 sec (Fig. 1). When the ISI was 0 sec, subjects heard the stimulus about 22 times before perceptual transition occurred (mean PTT was 7.5 sec), while when the ISI was 0.15 sec (mean PTT was 31.9 sec), they heard it about 62 times. Thus, subjects needed to hear more stimuli before perceptual transition when the ISI was 0.15 sec than when the ISI was 0 sec. One explanation for this phenomenon is satiation [4, 5, 2]. Subjects satiate when they hear repeated stimuli, which triggers a criterion shift in category boundary, which in turn leads to perceptual transition. According to the study by Efron [1], speech rate in everyday speech is ~12 phonemes/sec, (the length of one CV syllable is approximately 0.16 sec). Thus, an ISI of 0.15 sec is nearly equal in duration to one syllable. A pause equivalent to one syllable may delay satiation.

Perceptual transition occurred even in nonsense words. The mean PTT in the nonsense words, however, was longer than the mean PTT in the normal word /banana/ (Fig. 3). Moreover, the mean PTT varied according to word type. Specifically, the mean PTT for /nanaba/ was the longest in normal /banana/, emphasized /banana/, /bamana/, and /nanaba/. The nonsense word /nanaba/ transposes the order of the syllables in /banana/.

From these results, we may consider two main factors about VTE mechanisms. One is the masking effect. In general, the acoustic intensity of /ba/ is stronger than that of /na/. Moreover, in normal utterances, the acoustic intensity of a word-final syllable is weaker than that of a word-initial syllable (i.e., the ending of a word is pronounced weakly). The acoustic intensity difference between /ba/ and final /na/ in normal /banana/ (about 10.8 dB) was the same as that between initial /na/ and /ba/ in /nanaba/ (about 11.0 dB). Namely, the size of backward masking in /nanaba/ was almost the same as that in /banana/. This means that backward masking probably does not affect PTT. As for the case of emphasized /banana/, although the difference in acoustic intensity between /ba/ and the final /na/ is small (about 0.4 dB), as is seen in Fig. 1, the mean PTT in emphasized /banana/ was not so much longer than that in normal /banana/. Therefore, the difference in acoustic intensity probably does not affect perceptual transitions.

The other reason may be that /nanaba/ is a nonsense word. The processing of nonsense words in the brain may be different from meaningful words.

However, the mean PTT in the nonsense word /bamana/ did not significantly differ from that in the normal /banana/.

As for the illusory words, six subjects heard the word “banan”, when the ISI was 0 sec. When the ISI was 0.15 sec, however, subjects heard a range of illusory words. The perceived illusory words varied widely among the subjects when the ISI was 0.15 sec, which may be an effect of fading satiation. Alternatively, the mechanism underlying the occurrence of illusory words may differ from the case of ISI 0 sec. In any case, future experiments will be needed.

4. CONCLUSION

Although the verbal transformation effect has generally been thought to occur only with no pauses between repeated stimuli, we found that in this study it did occur in cases of inter stimulus interval of 0.15 sec. When the inter stimulus interval was 0 sec, 7.5 sec elapsed before perceptual transition occurred. However, when the inter stimulus interval was 0.15 sec, 31.9 sec elapsed. Thus, an increase in inter stimulus interval could have reduced satiation.

When the syllables in a real word were transposed, the perceptual transition time differed significantly. Differences in acoustic intensity between word-initial syllables and word-final syllables probably did not affect perceptual transition time, and we aslo found that perceived illusory words varied with inter stimulus interval.

Acknowledgment
This study is supported in part by a Grant-in-Aid for Scientific Research (No. 17K02707) from the Japan Society for the Promotion of Science. Our special thanks are due to Prof. Donna Erickson (Kanazawa Medical University) for her helpful advice.

5. REFERENCES

[1] Efron, R., 1963. Temporal perception, aphasia, and déjà vu. *Brain* 86, 403-424.
[2] Kondo, H. M. & Kashino, M. 2007. Neural mechanisms of auditory awareness underlying verbal transformations. *Neuroimage* 36, 123–130.
[3] Warren, R.M., 1961. Illusory changes of distinct speech upon repetition-The verbal transformation effect. *Br. J. Psychol*. 52, 249–258.
[4] Warren, R.M., 1968. Verbal transformation effect and auditory perceptual mechanisms. *Psychol. Bull*. 70, 261–270.
[5] Warren, R.M., 1985. Criterion shift rule and perceptual homeostasis. *Psychol. Rev*. 92, 574–584.
[6] Warren, R.M., Gregory, R.L., 1958. An auditory analogue of the visual reversible figure. *Am. J. Psychol*. 71, 612–613.