

The Effect of Different Latencies on Utterances of Japanese EFL Learners: A Psycholinguistic Study Based on Repeating and Open Question Tasks

MORISHITA, Miwa

Kobe Gakuin University

Abstract

Japanese EFL learners face many difficulties in terms of speaking. They often lack opportunities to speak English, and access to effective pedagogy of speaking has not been firmly established, with a limited amount of research on speaking due to its complex nature. In order to develop more effective pedagogical methods, the present study aims to shed light on the differences in speech processes between native English speakers and Japanese EFL learners at different proficiency levels. It examines the effect of different lengths of latencies and sentences on utterances of Japanese EFL learners as well as their use of vocabulary in comparison to native English speakers through a psycholinguistic experimental approach. The results demonstrated that the different lengths of latencies did not statistically affect either native English speakers or Japanese EFL learners, while the different lengths of sentences affected only the latter. Although the mean scores of tokens, types and word families differed markedly between native English speakers and Japanese EFL learners, the use of vocabulary was relatively similar between each group, with the preference of using vocabulary with a high rank of familiarity.

Key Words

speech processes, latency, automaticity, productive vocabulary

1. Introduction

Japanese EFL learners have had a reputation for being poor English speakers compared to their counterparts throughout the world. In fact, the mean score of native speakers of Japanese was the lowest in the world ranking in the speaking section of TOEFL iBT (Internet-based Testing) based on approximately 240,000 examinees who were tested between September 2005 and December 2006 (Educational Testing Service, 2007). It is said that the renewal of the TOEFL test, which now measures all four skills of English, has made it even harder for Japanese people to obtain a sufficient score to enter universities overseas, mainly because of speaking skill requirements. This can be regarded as a very serious situation in terms of international competitiveness, not only in the academic field, but also in business.

The purpose of the present study is to examine the differences in speech processes between native English speakers and Japanese EFL learners, who were given the different lengths of latencies (artificial pauses) and sentences in repeating and open question tasks.

The level of automaticity was investigated from the perspective of vocabulary, based on the idea that we need overall language knowledge (including grammatical knowledge) in order to produce words (utterances). Vocabulary used by each group was analyzed not only in quantitative but also qualitative approaches to find out how much and what kind of vocabulary is required for proper speech production based on utterances of native English speakers. The final goal is to explore psycholinguistically relevant speaking pedagogies intended for Japanese EFL learners.

2. Literature Review

2.1 Levelt's Spoken Language Processing Model

It is probably fair to say that most research on speaking has been based on the Levelt's spoken language processing model shown in Figure 1.

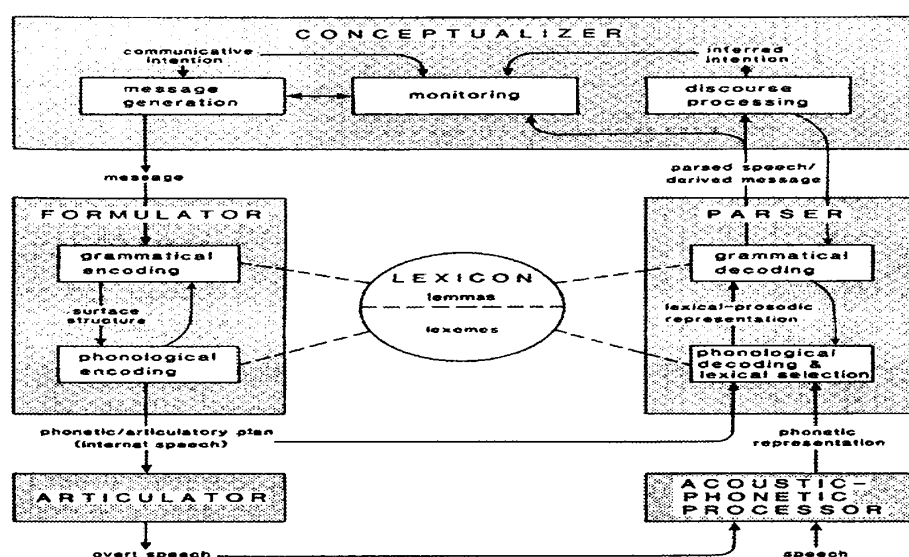


Figure 1. Schematic Representation of the Processing Components Involved in Spoken Language Use. Reprinted from Levelt, 1993, p.2.

Processes on the right represent listening and those on the left represent speaking. Speaking involves conceiving of an intention, selecting the relevant information to be expressed for the realization of this purpose, ordering this information for expression, and keeping track of what was said before. These activities are processed in what Levelt terms the *conceptualizer*. Then, in the *formulator*, preverbal messages are encoded in two steps: grammatical encoding and phonological encoding. The former accesses lemma information stored in one's mental lexicon and builds syntax. The latter accesses lexeme information to retrieve a phonetic or articulatory plan for each lemma and for utterance as a whole. The *articulator* retrieves successive chunks of internal speech from the buffer, where the phonetic plan is temporarily stored, and unfolds them for execution. Finally, a speaker has access to both his internal and overt speech in a self-monitoring stage back in the conceptualizer (Levelt, 1989, 1999).

Although this speech process is completed under time pressure, L1 speakers handle it in

parallel by way of automaticity, where they can comprehend what is being said, and think about what to say and how to say it at the same time without conscious attention to the linguistic code. In other words, “production has to be incremental, parallel and automatized in order to account for the enormous speed at which language is produced” (de Bot, 1992, p.6). Therefore, if speakers cannot perform the mental activities in real-time, both as listeners and as speakers, they will not be able to participate actively in communication (Versant with Ordinate® Technology, 2006).

2.2 Planning Time

Although the planning process is normally subconscious and highly automatic in L1 production, L2 speakers, especially EFL speakers, do not have ready-made plans (found in L1 production), which contribute to reducing the processing load (Mehnert, 1998). Yuan & Ellis (2003) investigated the effect of planning on L2 oral production. They gave participants an oral narrative task under three conditions: no planning, pre-task planning (10 minutes) and on-line planning (unlimited time given for the task), and measured fluency, grammatical / vocabulary complexity and accuracy of utterances. The result showed that pre-task planning enhanced grammatical complexity while on-line planning positively influenced grammatical complexity and accuracy. In addition, the pre-task planners were more fluent and used a greater variety of vocabulary than the on-line planners.

There are several other studies on planning time with mixed results (Foster and Skehan, 1996; Mehnert, 1998; Ortega, 1999). Among the reasons for the different results, especially in accuracy, may be the differences in task types, the time allocated to each task and/or definitions (measures) of variables (fluency, complexity and accuracy) in each study. Further research is required in this field, especially in EFL settings, because how much planning time should be given and how it should be presented in different tasks are essential information for oral English instruction, which is being gradually introduced in Japan in response to the governmental policy on the reform of English education (Ministry of Education, Culture, Sports, Science and Technology, 2003).

2.3 Research on the Speech Processes in EFL

Yokokawa, Yabuuchi, Suzuki and Morishita (2006) conducted one of the few research experiments on the speech processes in EFL in which utterances of 34 Japanese undergraduate students were analyzed using the *Versant for English*, a 10-minute computerized speaking test (Versant with Ordinate® Technology, 2006). This research intended to investigate the effect of a 3-week study abroad program on oral proficiency of the students based on a comparison in utterances for open questions between the pre-test and the post-test. The results showed that the mean response latency was greatly reduced in the post-test, indicating that students improved their language processing speed. It was also found that although the mean scores increased in all items of tokens, types and word families in the post-test, significant difference was only seen in tokens. This means that students did not necessarily build up their vocabulary *per se*. However, combined with the results of response latency, it is reasonable to assume that even a short-term study program helped students speed up access to and retrieval of vocabulary in both listening and speaking processes.

It was also found from the above data that only basic vocabulary with a high rank of familiarity (Yokokawa, 2006, in press) was used, half of which consisted of only about 20 words (Yokokawa, 2007). Considering that vocabulary familiarity can be a criterion for learnability, which is of some help to boost learners' motivation, it is an essential tool for deciding what kind of words should be taught in what order. It was also used in the present study.

2.4 Residual Problems

Although Yokokawa, Yabuuchi, Suzuki and Morishita (2006) has shown a possibility that the students who had participated in a short-term study program promoted automaticity in terms of information retrieval from the mental lexicon, how automaticity affected utterances was not fully examined. Therefore, it is necessary to control the time before speaking (by providing artificial pauses for planning) and examine the effect of different latencies on utterances. Since how different latencies affect what subsequent processes is also uncertain, the present study intends to investigate their effect on utterances in terms of simple repeating of sentences and spontaneous speech.

3. Experiment

3.1 Hypotheses

Three hypotheses have guided the present study.

- 1: The different lengths of latencies for repeating and open question tasks have an effect on the accurate reproduction and utterances of Japanese EFL learners.
- 2: The different lengths of sentences for repeating tasks have an effect on the accurate reproduction by Japanese EFL learners.
- 3: Japanese EFL learners' proficiency affects utterances in terms of vocabulary in open question tasks, and they differ greatly from those of native English speakers.

3.2 Methods

Participants

36 Japanese undergraduate and postgraduate students learning English as a foreign language and 15 native English speakers living in Japan participated in the experiment. None of them had auditory or visual disorders. The Japanese EFL students were divided into high and low English proficiency groups according to the scores of the Versant for English. The score range of the high proficiency group was 40 to 57 and that of the low proficiency group was 24 to 39 (Full score = 80).

Tasks

In the experiment, two types of tasks – six sentences for repeating tasks (Set A) and three questions for open question tasks (Set B) – were prepared as shown in Table 1.

Table 1. Experimental Sentences and Questions

Set A	1 War broke out suddenly.
	2 Leave town on the next train.
	3 It's supposed to rain hard tomorrow, isn't it?
	4 Traffic is a huge problem in California.
	5 They play loud music all through the night when he is trying to sleep.
	6 There are three basic ways where a story might be told to someone.
Set B	1 Do you prefer to work by yourself or in a group? Please explain.
	2 Do you prefer to communicate by telephone or by writing letters? Why?
	3 Do you think television has had a positive or negative effect on family life? Please explain.

Although repeating tasks are rarely used for research on planning time, they were used in the present study in order to compare the different effects of latencies on repeating tasks and open question tasks. In both tasks, there were three different lengths of latencies and each condition had two sentences in Set A and one question in Set B as shown in Table 2.

Table 2. Latency Conditions for Set A and Set B

Set A	Repeating tasks	No latency	2
		Short latency (5 seconds)	2
		Long latency (10 seconds)	2
Set B	Open question tasks	No latency	1
		Short latency (10 seconds)	1
		Long latency (20 seconds)	1
Total			9

Unlike the case of Set B, latencies in Set A were not for planning but mainly for articulatory rehearsal, since auditory memory is believed to disappear within a few seconds without being rehearsed in the phonological loop in working memory (Baddeley, 1986, 2000). This means that, 0 condition requires no rehearsal, while 5 and 10 second conditions require it for accurate reproduction.

In addition, in Set A, there were three different sentence lengths as shown in Table 3.

Table 3. Sentence Conditions for Set A

			# of words	# of syllables
Set A	Repeating tasks	2 short sentences	4	6
			6	6
		2 medium sentences	8	11
			7	11
		2 long sentences	14	16
			13	16
Total		6 sentences		

The lengths of sentences were not based on the number of words but on the number of syllables with reference to Kohno (1993). He defined seven plus or minus two syllables, whose intervals are less than about 330 ms, as a Perceptual Sense Unit (PSU), a perceptible unit of human memory. Therefore, tasks are considered to be increasingly difficult to process for Japanese EFL learners as sentences become longer. In contrast, native English speakers, who are used to holistic processing in PSUs, can generally repeat sentences that contain more than seven syllables because they are very familiar with words, phrase structures, and other common syntactic forms. Thus, “if a person habitually processes five-word phrases as a unit (e.g. “her really big apple tree”), then that person can usually repeat utterances of 15 or 20 words in length” (Versant with Ordinate® Technology, 2006, p.2). This is what Miller (1956) referred to as “chunking,” a grouping process of individual information into larger units.

All sentences and questions were recorded by a native speaker of English. For Set A, beeps were inserted after each sentence with three types of latencies (0, 5 and 10 seconds), and after 10 seconds of the response time, a sentence was subsequently given. For Set B, after each question was repeated twice, the first beeps were inserted with three types of latencies (0, 10 and 20 seconds) and the second beeps were inserted after 20 seconds to give notification of the end of the response time.

Procedures

Each participant in the experiment was tested individually while seated next to the experimenter. Before the experiment started, the instructions were given and the participants were required to sign a letter of consent. Then, the participants wore headphones and a miniature microphone and listened to a set of practice tasks. After each beep, latencies of the different lengths were provided, and participants were required to repeat three sentences with different latencies for Set A and answer an open question for Set B.

After the practice session, they were able to ask questions if needed and when they were prepared, the main tasks started. The main tasks were conducted in the same manner as the practice session on a laptop personal computer running Microsoft Windows XP. Their utterances were recorded with an IC recorder and transferred to the computer. At the end of the experiment, the participants were required to answer a brief questionnaire about their performances. The whole process took approximately 30 minutes per individual.

After the experiment, all the utterances were transcribed in an orthographic manner, not based on phonetics. Minor pronunciation errors were allowed, but words pronounced in a way where participants obviously did not know them were not counted. In addition, all kinds of fillers indicating hesitation were transcribed as “uh.” All transcription was then rechecked by a native speaker of English and used for vocabulary analyses from different perspectives, including the mean scores of tokens, types and word families as well as frequency based on Paul Nation’s RANGE 32GSL (2005) and familiarity based on Yokokawa (2006, in press).

4. Results and Discussion

Data were analyzed by one-way ANOVA with a significance level of .05 on the scores of accurately repeated words, tokens, types and word families as well as response latencies as dependent variables, and the lengths of latencies and the lengths of sentences as independent variables.

4.1 Results 1: Mean Scores of Utterances and Response Latencies in the Different Lengths of Latencies

The mean scores of accurately repeated words in Set A, those of tokens, types and word families as well as response latencies in Set B, all in the different lengths of latencies, are shown in Table 4, 5, and 6 respectively.

Table 4. Mean Scores of Accurately Repeated Words and SD in the Different Lengths of Latencies in Set A

	0	5	10	<i>p</i>
NS	9.68 (0.73)	9.64 (1.24)	9.88 (0.42)	
Upper	6.67 (2.34)	7.50 (2.14)	7.34 (2.76)	
Lower	5.49 (2.51)	5.81 (2.50)	5.00 (2.43)	

Note. All scores were calculated on a scale of 0 to 10.

The different lengths of latencies had no significant effect on the mean scores of accurately repeated words in Set A for either Japanese EFL learners or native English speakers, with a slight increase, however, in the case of native English speakers and upper level students. In contrast, there were statistically significant differences between each group in all conditions except between upper / lower level students in 0 second condition according to the multiple comparison. This means that upper level students' utterances tended to be almost the same as those of lower level students when there was no latency but increase when latencies were provided. Therefore, it might be assumed that the former could effectively rehearse sentences in the phonological loop.

Table 5. Mean Scores of Tokens, Types & Word Families and SD in the Different Lengths of Latencies in Set B

		Tokens	Types	Families	<i>p</i>
NS	0	55.2 (10.90)	39.3 (5.60)	34.1 (4.84)	
	10	56.9 (7.78)	40.9 (5.54)	35.3 (4.95)	
	20	61.5 (8.67)	43.0 (4.86)	38.2 (3.83)	
Upper	0	26.5 (10.14)	17.8 (5.82)	15.8 (5.17)	
	10	27.8 (9.62)	19.4 (5.98)	17.4 (5.48)	
	20	30.1 (6.98)	19.3 (5.30)	17.5 (5.00)	
Lower	0	17.5 (7.37)	12.4 (4.32)	10.6 (3.84)	
	10	16.4 (9.70)	12.4 (6.00)	11.0 (5.59)	
	20	18.4 (7.59)	13.0 (4.44)	11.5 (3.99)	

The different lengths of latencies had no significant effect on the mean scores of types, tokens and word families in Set B for either Japanese EFL learners or native English speakers. However, all tokens, types and word families were increased in the case of native English speakers and upper level students, with the former demonstrating a larger increase, indicating that if latencies become even longer, a significant effect may be observed. No such

shift was shown in the case of lower level students.

Table 6. Mean Response Latencies (second) and SD in the Different Lengths of Latencies in Set B

	0	10	20	<i>p</i>
NS	0.73 (0.27)	0.76 (0.18)	0.73 (0.28)	
Upper	0.85 (0.34)	0.88 (0.35)	0.93 (0.59)	
Lower	1.22 (0.49)	1.22 (0.60)	1.07 (0.33)	

The different lengths of latencies had no significant effect on response latencies in Set B for either Japanese EFL learners or native English speakers. The results of the multiple comparison, however, showed significant differences between each group as follows: [1] between native English speakers / lower level students ($F(1, 33) = 11.89, p < .01$) and upper / lower level students ($F(1, 35) = 6.46, p < .05$) in 0 second condition, [2] between native English speakers / lower level students ($F(1, 31) = 7.63, p < .01$) in 10 second condition. Therefore, although no significant difference appeared within each group, there was a tendency among lower level students to respond quicker when latencies became longer.

4.2 Results 2: Mean Scores of Accurately Repeated Words in the Different Lengths of Sentences

Table 7. Mean Scores of Accurately Repeated Words and SD in the Different Lengths of Sentences in Set A

	Short	Medium	Long	<i>p</i>
NS	9.94 (0.30)	9.79 (1.12)	9.46 (0.89)	
Upper	8.58 (1.79)	7.75 (2.26)	5.18 (1.85)	**
Lower	6.95 (2.10)	5.85 (2.33)	3.50 (1.66)	**

Note. All scores were calculated on a scale of 0 to 10.

There was no significant difference in the mean scores of accurately repeated words in the different lengths of sentences internally in native English speakers, as shown in Table 7, while significant differences were found in all conditions in both upper level students ($F(2, 101) = 26.39, p < .01$) and lower level students ($F(2, 113) = 27.41, p < .01$) except between short-length and medium-length sentences in upper level students. These results mean that the different lengths of sentences did not affect native English speakers, while Japanese EFL learners tended to have difficulty in repeating when sentences were long. According to the findings that upper level students showed no significant difference in repeating short-length and medium-length sentences, it would be reasonable to assume that lower level students could only manage to memorize the short-length sentences (6 syllables), which can be perceived as a PSU, while upper level students were able to reconstruct the medium-length sentences (11 syllables), utilizing their grammatical and/or lexical knowledge.

4.3 Results 3: Vocabulary Analyses

Table 8. Mean Total Scores of Tokens, Types and Word Families in Set B

	Tokens	Types	Families
NS	173.6	33.1	23.3
Upper	84.5	13.4	10.3
Lower	52.3	9.0	6.8

The mean total scores of tokens, types and word families of native English speakers were more than double those of upper level students and more than three times as many as those of lower level students as shown in Table 8. These scores correspond to words per minute (wpm), since there were a total of three answers to 20 second-responses, which is equivalent to one minute per participant. The scores of the native English speakers' utterances were in line with the general opinion that their natural speech speed is around 180 wpm. In contrast, the low scores of Japanese EFL learners reflected, with their slow processing speed of fewer than 100 wpm, that they may have a difficulty in listening to English spoken at natural speed, let alone speaking.

Unlike the major differences in the mean total scores of tokens, types and word families between native English speakers and Japanese EFL learners, vocabulary familiarity was, on the whole, relatively similar between each group. The data based on the lists of visual and audio vocabulary familiarity (Yokokawa, 2006, in press) in Set B are shown in Table 9, 10, 11 and 12.

Table 9. Occurrences of Lemmas and Their Visual and Audio Familiarity in Set B

	NS			Upper			Lower		
%	Lemmas	Visual	Audio	Lemmas	Visual	Audio	Lemmas	Visual	Audio
0-50%	27	6.51	5.43	16	6.43	5.33	12	6.44	5.43
50-60%	14	6.42	5.37	8	6.53	5.37	7	6.49	5.28
60-70%	22	6.21	5.50	11	6.39	5.54	9	6.25	5.28
70-80%	45	6.27	5.41	18	6.34	5.37	15	6.45	5.83
80-90%	95	6.11	5.51	37	6.36	5.66	27	6.27	5.31
90-100%	233	5.47	5.09	112	5.95	5.29	80	6.07	5.41
Total	436	5.86	5.28	202	6.15	5.39	150	6.21	5.42

The scores of lemmas used in about half of the utterances were 27, 16, and 12 in native English speakers, upper and lower level students respectively as shown in Table 9. The total rates of their mean visual familiarity were 5.86, 6.15 and 6.21 and those of audio familiarity were 5.28, 5.39 and 5.42. Overall, the mean rate of visual familiarity was much higher than that of audio familiarity in each group with more than 0.5 point difference, and the former gradually decreased as occurrences increased, while such shift was not seen in the case of the latter. A possible reason for the fact that the former is more related to speech production than the latter is that participants may have visualized written forms (letters) during speech planning (Yokokawa, 2007).

Table 10. Occurrences of Lemmas based on Visual Familiarity Ranks in Set B

Familiarity (points)	NS	%	Upper	%	Lower	%
6.0 - 6.9	215	49.31	146	72.28	107	71.33
5.0 - 5.9	119	27.29	33	16.34	27	18.00
4.0 - 4.9	39	8.95	11	5.44	4	2.67
3.0 - 3.9	14	3.21	4	1.98	1	0.67
2.0 - 2.9	1	0.23	0	0	1	0.67
Not in the list	48	11.01	8	3.96	10	6.66
Total # of lemmas	436	100.00	202	100.00	150	100.00

In addition, about half of the lemmas used by native English speakers were in the 6.0 and 6.9 point ranges in visual familiarity as shown in Table 10, meaning that even native English speakers used a limited number of familiar words out of 3000 high-frequency words in the BNC (British National Corpus) frequency list. However, considering that native English speakers had a much higher rate of lemmas (27.29%) in the 5.0 and 5.9 point ranges than that of Japanese EFL learners, increasing the number of basic vocabulary for production from the beginning may be a shortcut to becoming a fluent speaker as well as a proficient listener.

Table 11. List of Vocabulary Frequency based on the Top 30 Words

	NS	Upper	Lower
1	I	I	uh
2	to	uh	I
3	uh	because	because
4	the	can	to
5	and	to	is
6	a	the	<i>television</i>
7	you	we	and
8	it	<i>think</i>	<i>prefer</i>
9	is	<i>like</i>	<i>telephone</i>
10	because	<i>prefer</i>	with
11	by	and	can
12	<i>prefer</i>	<i>television</i>	<i>family</i>
13	<i>work</i>	so	<i>think</i>
14	that	<i>work</i>	<i>communicate</i>
15	do	a	in
16	of	is	the
17	on	by	<i>work</i>
18	can	if	it
19	so	do	a
20	<i>like</i>	<i>family</i>	<i>effect</i>
21	<i>people</i>	not	<i>like</i>
22	<i>television</i>	<i>communicate</i>	by
23	my	<i>group</i>	myself
24	not	<i>positive</i>	<i>positive</i>
25	in	<i>telephone</i>	<i>group</i>
26	<i>think</i>	with	<i>negative</i>
27	more	in	so
28	myself	it	has
29	or	other	<i>want</i>
30	at	on	have

Note. Content words are in italics.

According to Table 11, the most obvious difference in vocabulary frequency between each group is that the use of articles (“a” and “the”) decreased, with native English speakers using the most and lower level students using the least. This result reflects the fact that native speakers of Japanese are notoriously weak in the use of articles, which do not exist in Japanese, and in fact, even in repeating tasks, Japanese EFL learners tended to omit articles. In addition, the ranks of the words “think” and “because” were higher in the case of Japanese EFL learners than native English speakers, reflecting the fact that the former tended to use the phrase “I think ... because ...” even when they had no particular reason and could not continue utterances.

Table 12. Characteristics of Vocabulary based on the Top 30 Words

Common words to all groups (18)
<i>like, prefer, television, think, work</i> , a, and, because, by, can, I, in, is, it, so, the, to, uh
Common words between native English speakers and upper level students (3)
do, not, on
Common words between upper level students and lower level students (6)
<i>communicate, family, group, positive, telephone</i> , with
Common word between native English speakers and lower level students (1)
myself
Words only found in native English speakers (8)
<i>people</i> , at, more, my, of, or, that, you
Words only found in upper level students (3)
if, other, we
Words only found in lower level students (5)
<i>effect, negative, want</i> , has, have
<i>Note.</i> Content words are in italics.

If we closely look at the characteristics of vocabulary used by each group shown in Table 12, there are some more interesting features as follows: [1] all groups shared a relatively large number of basic words (18 out of 30), [2] native English speakers and upper level students shared the words for negative expressions (“do” and “not”), where they needed to explain the basic reasons for their negative opinions, [3] Japanese EFL students, especially lower level students, used a large number of content words, most of which had appeared in the questions, [4] native English speakers used more function words to make longer sentences, [5] native English speakers used “you” or “people” and upper level students used “we” when they talked in generalities, while lower level students did not use these pronouns and [6] upper level students used “if” in an effort to make complex sentences such as conditional or subjective clauses.

In summary, the results above indicate that although lower level students had almost nothing in common with native English speakers, upper level students had much in common with them. Even though upper level students used significantly fewer words than native English speakers (see Table 8), both groups actually uttered relatively similar vocabulary as shown in Table 11 and 12. This means that upper level students lacked processing speed of vocabulary rather than vocabulary knowledge, indicating a possibility, together with the results in Table 5, that they will both increase processing speed and utter more high-frequency words, which they know but cannot use as productive vocabulary, if longer latencies are provided.

4.4 Results from the Questionnaire

The results from the questionnaire about their performances enable us to infer what participants may have been thinking during each task. It is suggested, especially in the case of Japanese EFL learners, that although the different lengths of latencies psychologically affected them, they were not necessarily reflected in their utterances. The following is a summary of the results from the questionnaire of each group.

Native English speakers

- The different lengths of sentences in Set A did not affect them at all.
- Shorter latencies were better in Set A, because they could repeat sentences as soon as possible without rehearsing them in their heads.
- Shorter latencies were better in Set B, because longer latencies forced them to think about more things than they could speak within 20 seconds.

Upper level students

- They tried to understand the meaning of sentences even in the repeating tasks in Set A and if they understood them in Japanese, they successfully repeated them.
- They tended to grasp the main points and paraphrase sentences in Set A.
- Some of them preferred longer latencies for long sentences in Set A, because they could think about the sentence structure.
- They tended to think about their answers in Japanese and translate them into English in Set B.

Lower level students

- They tried to repeat exactly what they heard without thinking about the meanings in Set A.
- They had difficulty in listening especially in Set A.
- They could not find their answers even in Japanese in Set B.
- They took time to find the words and phrases that they could use in English in Set B.
- Some of them preferred shorter latencies in Set B, because they could use the words and phrases which had appeared in the questions for their speech.

According to the results above, native English speakers found it bothersome, even if they accurately repeated sentences, to hold verbal information using a temporary store and an articulatory rehearsal system in the phonological loop when latencies were long. Their utterances slightly increased, however, as latencies became longer (see Table 4), implying that articulatory rehearsal might have had a positive effect on utterances.

Upper level students seem to have experienced higher cognitive load and executed many more different types of processing compared to lower level students. Considering that the results of the experiment showed that the former tended to increase utterances when latencies were provided (see Table 4 and 5), they may have utilized these latencies for both articulatory rehearsal and planning. Meanwhile, they tried to translate their opinions created in Japanese into English, perhaps indicating the adverse effect of the grammar translation method which still predominates English classes in Japan. This phenomenon may also indicate that they had both L1 and L2 systems in the formulator (de Bot, 1992). De Bot pointed out that “a concept has to be expressed in a particular language which does not have the lexical items needed to express that concept, or for which the relevant item cannot be found (in time)” (1992, p.8) is a problem that often faces non-balanced bilinguals.

Lower level students did not seemingly utilize the provided latencies for either articulatory rehearsal or planning. They rather preferred no latency conditions because they could immediately repeat sentences in repeating tasks or use words and phrases which had appeared in questions for answering them in open question tasks.

Thus, even if the results did not support the effect of the different lengths of latencies in

any statistically significant manner, both native English speakers and Japanese EFL learners may have perceived the differences and reacted to them in diverse ways.

4.5 Hypotheses Verification

The findings in 4.1 did not verify Hypothesis 1: *The different lengths of latencies for repeating and open question tasks have an effect on the accurate reproduction and utterances of Japanese EFL learners*; however, with respect to response latencies, a minor effect of the different lengths of latencies was shown among lower level students.

The findings in 4.2 verified Hypothesis 2: *The different lengths of sentences for repeating tasks have an effect on the accurate reproduction by Japanese EFL learners*.

The findings in 4.3 verified Hypothesis 3: *Japanese EFL learners' proficiency affects utterances in terms of vocabulary in open question tasks, and they differ greatly from those of native English speakers*; however it only applies to the mean scores of tokens, types and word families. In terms of vocabulary familiarity, all groups preferred to use words that they were comfortable with, i.e. high-ranking words in the familiarity lists, with minor differences between them.

4.6 Pedagogical Implications

The results from the experiment showed that there was no effect of the different lengths of latencies on utterances of native English speakers, who might have integrated automaticity in their speech processes, while Japanese EFL learners showed mixed results.

First of all, the different lengths of latencies showed little or no effect on utterances of both native English speakers and Japanese EFL learners, but this was seemingly not because of automaticity in the case of the latter in view of previous research. One of the possible explanations is that Japanese EFL learners, especially lower level students, could not properly comprehend what was said or asked in English. This implies the need for teaching listening in order to improve speaking. Furthermore, since latencies of 10 or 20 seconds were not long enough for Japanese EFL learners, they could not effectively utilize the time for planning in Set B. In this regard, guidance on how to use the different lengths of planning time, which has been pedagogically overlooked so far, is surely required. In view of the fact that giving planning time does not always have beneficial results, as observed in several studies including the present one, a decision on how much planning time should be given to what kind of tasks and for what level of students, which has generally been made by teachers' instinct alone, is also of critical importance.

In contrast to the case of latencies, the different lengths of sentences had a great effect on the accuracy of repeating tasks for Japanese EFL learners. Unlike other similar tasks such as oral reading and shadowing, which mainly act on the articulator in the Levelt's spoken language processing model (see Figure 1), repeating tasks require processing functions such as grammar encoding and phonological encoding in the formulator (Kadota, 2007). In fact, lower level students could not accurately repeat the medium-length and long-length sentences (exceeding a PSU), indicating they could not properly access processing capacity due in part to lack of grammatical and lexical knowledge. Given repeating tasks are not mechanical tasks in reality but cognitively demanding ones, where structural understanding is required, we

should have another look at them. Being able to be implemented in nearly any classroom environment, they are especially useful for relatively large classes in Japan.

In terms of vocabulary, both limited exposure and lack of practice may hinder the successful passage of words from receptive to productive vocabulary (Laufer and Paribakht, 1998) for Japanese EFL learners. However, even in the case of most native English speakers, there are a large number of low-frequency words for which they only have receptive knowledge. Given the fact that native English speakers used slightly less familiar words compared to Japanese EFL learners both in visual and audio familiarity, Japanese EFL learners can improve their speaking by extending the range of their spoken vocabulary focusing on basic words. Thus, it might be a good idea for them to identify receptive vocabulary that can be and should be used for production by referring to vocabulary familiarity and concentrate on mastering that.

5. Concluding Remarks and Further Studies

Although a variety of implications were obtained from the present study, several questions remain to be discussed. One of the main reasons why we did not observe any statistically significant effect of the different lengths of latencies might be related to the unprecedented nature of the present study. This was a totally new approach and the experimenter herself needed to decide every detail, such as lengths of latencies and sentences. Thus, further trial and error experiments may be necessary to obtain more accurate results.

In addition, in the present study, the mean scores of the Versant for English of the high proficiency group and the low proficiency group were 44.29 and 34.32 respectively. Although these groups were divided on the score of 40, most of the students were actually at a similar level, which may have affected the results.

Another problem was that, in repeating tasks, there were some cases where the participants obtained a perfect score even if they included unnecessary words in sentences or repeated words in the wrong order, because they were not marked as a negative. In contrast, although only the scores of accurately repeated words were counted, upper level students tried to paraphrase the original sentences as ones with the same or similar meanings rather than simply repeating them. This is a type of speaking strategy that is especially useful in conversation and when learners master it, “they are able to operate effectively with a small productive vocabulary” (Nation, 1990, p.110). In this regard, the scores of accurately repeated words may not necessarily have reflected their language performance *per se*.

Although the present study investigated only fluency by the scores of tokens, types and word families, other variables such as complexity and accuracy could also have been investigated in some way. Since upper level students used quite a few function words as shown in Table 11 and 12, they may have uttered more grammatically complex sentences than lower level students. Moreover, although the relationship between visual and audio familiarity was not sufficiently examined in the present study, further research in this field is also needed to find out what types of input are related to what output activities.

In conclusion, the effect of latencies on utterances is a highly important issue for Japanese EFL learners to overcome their weakness in speaking, especially in terms of processing speed, and needs further investigation. To this end, more studies, not only on this topic but

also on speaking as a whole, should be urgently implemented in EFL settings. Considering that the four skills of English are all important and interrelated, research in which more than two language skills are involved, such as speaking and listening from the perspective of sound, or speaking and writing from that of production, is particularly desirable.

References

- Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4 (11), 417–423.
- de Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics*, 13 (1), 1–24.
- Educational Testing Service. (2007). Test and score data summary for TOEFL® internet-based test. Retrieved December 18, 2007, from <http://www.ets.org/Media/Research/pdf/TOEFL-SUM-0506-iBT.pdf>
- Foster, P. & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18 (3), 299–323.
- Kadota, S. (2007). Science of shadowing and oral reading. Tokyo: Cosmopier.
- Kohno, M. (1993). Perceptual sense unit and echoic memory. *International Journal of Psycholinguistics*, 9 (1), 13–31.
- Laufer, B. & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48 (3), 365–391.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1993). The architecture of normal spoken language use. In G. Blanken, J. Dittman, H. Grimm, J. C. Marshall and C. Wallesch (Eds.), *Linguistic disorders and pathologies: An International handbook* (pp.1–15). Berlin: Walter de Gruyter.
- Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83–122). Oxford: Oxford University Press.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20 (1), 83–108.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63 (2), 81–97.
- Ministry of Education, Culture, Sports, Science and Technology. (2003). Action plan to cultivate "Japanese with English abilities." Retrieved December 18, 2007, from http://www.mext.go.jp/b_menu/houdou/15/03/03033101/001.pdf
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.
- Nation, I. S. P. (2005). RANGE and FREQUENCY programs for Windows based PCs. Retrieved December 18, 2007, from <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx#top>
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21 (1), 109–148.
- Versant with Ordinate® Technology. (2006). *Versant for English technical manual*. Texas: Harcourt Assessment, Inc.

- Yokokawa, H. (Ed.). (2006). Database for second language pedagogy and research: English vocabulary familiarity of Japanese EFL learners <Visual version>. Tokyo: Kuroshio Publishers.
- Yokokawa, H., Yabuuchi, S., Suzuki, M. & Morishita, M. (2006). Vocabulary and response latencies in the spontaneous speech of Japanese EFL learners: An analysis of data from a short-term overseas study tour. *IEICE Technical Report, TL2006-33*, 13-18. Tokyo: The Institute of Electronics, Information and Communication Engineers.
- Yokokawa, H. (2007). Frequency and familiarity of vocabulary in the spontaneous speech of Japanese EFL learners. *Journal of the School of Languages and Communication, Kobe University*, 3, 59-77.
- Yokokawa, H. (Ed.). (in press). Database for second language pedagogy and research: English vocabulary familiarity of Japanese EFL learners <Audio version>.
- Yuan, F. & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24 (1), 1-27.