L2 Japanese Proficiency and Working Memory Capacity*

Yuki Hattori
The Ohio State University

Abstract
This paper presents a newly created L2 Japanese Speaking Span Test (L2JSST) and reports preliminary results on its validity based on ten English-speaking learners of Japanese. Intermediate-high to advanced Japanese learners took two Japanese language proficiency tests, the L2JSST, and the L1 English Speaking Span Test (L1ESST). The results show that individuals with higher L1ESST scores tend to have higher L2JSST scores. Language proficiency also correlates with the L2JSST scores. Although a high accuracy rate was overserved in produced sentences from L2JSST, when prosody is considered, some prosodic errors were noticeable, such as a low success rate of producing proper downsteps, and pronunciation errors. As the prosodic difficulty in L2 Japanese was reported by Goss and Nakayama (2011), the acquisition of prosody seems challenging for speakers of intermediate to advanced Japanese.

Key words
Working memory, L2 Japanese speaking span test, English speaking span test, Japanese proficiency test, Minimum Japanese test, prosody

* I am grateful to the participants of the study. I would like to thank Dr. Mineharu Nakayama for his assistance in the project. Of course, all shortcomings are mine. This research has been partially supported by the Nissen Chemitec of America Scholarship by The Ohio State University Department of East Asian Languages and Literatures. Their support is gratefully acknowledged.

1. Introduction

The effect of Working Memory (WM) on L2 language skills is a contentious topic in psycholinguistics. WM is generally understood as memory for the temporary holding and manipulating of information during the performance of cognitive tasks such as language comprehension, reasoning, and learning (Baddeley 1986; 2000; Baddeley and Hitch 1974). Baddeley (1986) claims that WM plays an important role in language processing. It is normally considered that the capacity of WM differs individually, and learners with larger Working Memory Capacity (WMC) store and process information more effectively than those with smaller WMC. The field of L2 speech production and WMC is still young and developing. Studies have yielded mixed results concerning the correlation between L2 speech production and WMC (i.g., Fortkamp 1999; 2000; Mota 2003, etc.). Daneman and Green (1986) created the Speaking Span Test (SST) as a tool to measure WMC in English. To our knowledge, however, there is no Japanese SST. In the current study, WMC scores measured by L1 English SST (L1ESST) will be compared with those measured by the newly created L2JSST. The L2JSST scores are also compared with the scores of two Japanese language tests, the PT (Japanese Language Proficiency Test) (Itomitsu and Nakayama 2005) and the MJT (Minimal Japanese Test) (Maki, Dunton, and Obringer 1999), to determine the validity of the L2JSST as a tool to measure WMC. Although sentence complexity and prosodic errors do not affect L2JSST scores, produced sentences in L2JSST will be observed to confirm the difficulty of prosodic acquisitions as many researchers claim (Goss and Nakayama 2011; Ayusawa 2003).

2. L2 Speech production and working memory capacity

Mixed results have been shown regarding the correlation between L2 speech production and WMC. Researchers including Fortkamp (1999, 2000), Weisshemier and Mota (2009), and Mizera (2006) used the SST to measure WMC for learners of L2 English. Most of these studies found that there was a correlation between WMC and oral fluency. Fortkamp (1999, 2000) used the SST to measure WMC for Portuguese-speaking learners of L2 English. She found that there was a correlation between WMC and lexical density and oral fluency. Note that oral fluency in her study included accuracy, complexity (producing a larger number of dependent clauses), pauses, and hesitations.

Weisshemier and Mota (2009) also tested Portuguese-speaking learners of L2 English and their WMC. They measured accuracy, fluency, and complexity, and found a correlation between the WMC and the oral proficiency level, i.e., learners with a larger WMC produced sentences faster and made fewer errors. Based on the results, Weisshemier and Mota suggest that learners in the larger WMC group had a greater number of cognitive resources to obtain new pieces of information or employ speaking strategies. Mota (2003) tested ESL learners whose L1 languages were Chinese, German, Hebrew, Indonesian, Japanese, Korean, Portuguese, and Turkish, and also observed that there was a positive relationship between WMC and L2 fluency, including speech rate, pauses, accuracy, and hesitations.

Mizera (2006), on the other hand, has a different result from the researchers mentioned above. They carried out an experiment for Spanish-speaking learners of L2 English using the SST, the math span test, and a non-word repetition test to measure WMC. They did not find a strong correlation between an individual's WMC and L2 oral fluency. This is not surprising because fluency includes so many different factors and it is difficult to determine whether the speaker is fluent in the target language or not. It is important to note that having a smaller WMC does not necessarily mean that learners cannot speak the language.

Given these previous studies, the field of L2 speech production and WMC is still developing.

However, many scholars believe that WM affects L2 speech production. Therefore, it can be assumed that L2 Japanese learners with larger WMC (i.e., higher span test scores) are more skilled in using their WMC effectively. Since there have been no studies conducted using a JSST, the L2JSST was created based on Daneman and Green's (1986) L1ESST and Fortkamp's (2000) L2ESST. Having a valid and reliable L2JSST is the first step before investigating the correlation between WMC and L2 Japanese fluency.

3. L2 Japanese speaking span test

3.1 Procedure

Fortkamp's (2000) L2ESST, which was modeled after Daneman and Green's (1986) original English SST, was used to make this study's L2JSST. In Fortkamp's test, there is a total of three trials and 60 unrelated words. Each trial has two-, three-, four-, five-, and six-word sets.

A total of 70 unrelated words are used in our L2JSST. There are five sets of two, three, four, and five words. Each set includes two, three, four, and five-word sets; referred to as a trial. The entire test has five trials. The words are all nouns with three and four morae, chosen from Japanese textbooks. Learners are presented these words on their computer screens, displayed one at a time for ten milliseconds. The next word appears in the same position where the previous word appears. This procedure was repeated until the letter "Q" was shown. Their task was to read each word silently and use the words they saw in generating sentences. Each word should be included in one and only one complete sentence. Hence, for one set, learners would produce 14 different sentences. The target words in the sentences should be in the same forms, and the order of the sentences must follow the order of the words as shown on the screen. The sentences produced do not have any length restrictions, but they should be syntactically correct and semantically coherent in Japanese. An example of the test is illustrated in Figure 1. A sample list of the words can be found in the Appendix. For instance, learners are presented with a set of words, one by one: 掃除 (*sooji*, 'cleaning'), ズボン (*zubon*, 'pants'), 歴史 (*rekishi*, 'history'). Then, they must produce sentences with those words such as ; (i) 昨日、部屋の掃除をした (*kinoo, heya-no sooji-o shita* 'Yesterday, I cleaned my room'), (ii)新宿でズボンを買った (*Shinjuku-de zubon-o katta* 'I purchased a pair of pants in Shinjuku'), (iii)私の専攻は歴史だ (*Watashi-no senkoo-wa rekishida* 'My major is history.').

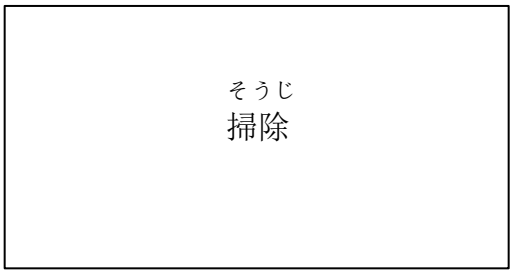

Figure 1: Speaking span test example

3.2 Measurement

Two types of scores were given to the participants to measure their WMC: strict and lenient speaking span scores (Daneman and Green 1986; Daneman 1991; Weissheimer and Mota 2009). A strict speaking span score was used to count the total number of sentences participants produced. In this score, sentences should include the same word form and order as shown on the screen. A

lenient speaking span score was also used to determine the number of sentences participants generate, but the word form and word order in the sentence can be different (i.e., 練習 *renshuu* 'practice' → 練習する*renshuu-suru* 'to practice'). Fourteen credits were given to the learner when they complete an entire trial with syntactically and semantically correct sentences. The grammatically accurate production of all five trials (70 words) would yield 70 credits/points, meaning that the highest possible score is 70 in this test, the total and the maximum number of words each learner will recall.

4. Experiment

Eleven L1 English-speaking learners of Japanese participated in this study and were nominally compensated. However, one participant's data was not included in the final analysis due to a technical issue. When the L2JSST was conducted, the participants were students in the 4th and 5th levels of Japanese language classes at a Midwestern US university. Learners were given two Japanese language proficiency tests: the PT (Itomitsu and Nakayama 2005) and MJT (Maki, Dunton and Obringer 1995). The PT consists of three sections: *Kanji* reading (16 questions, 4 minutes maximum), Vocabulary (16 questions, 8 minutes maximum), and Reading (8 questions, 18 minutes maximum). It is a multiple-choice test. The PT was created based on published sample questions of Levels 2 (N2) and 3 (N3) Japanese Language Proficiency Tests. The MJT has two versions: one was used in a practice round and the other for the test trial, where the individual listened to the audio and filled in blanks using *hiragana*, the Japanese phonetic alphabet. All skills except speech production were examined with the two tests. Following the tests, the L2JSST and the L1ESST (Fortkamp 2000) were administered. Because the entire procedure took about two hours in total, it was divided into two parts (two days). Day 1 consisted of the signing of a consent form, the administration of a background questionnaire, the PT, and MJT. The two SST were conducted on Day 2. The two sessions were carried out online (via Zoom) remotely each day and lasted approximately one hour.

5. Results

A summary of the test scores is listed in Table 1. The results show that the average ESST score (42.4/60=70.6/100) was higher than the average JSST score (37.9/70=54.1/100). A regression analysis shows that there is a highly significant correlation between the ESST and the JSST scores (Multiple r=0.81). The individuals tend to find L2SST more taxing than L1SST. There is a positive correlation between Japanese proficiency level (total scores of PT and MJT=JLPT) and JSST scores (Multiple r=0.64), indicating that those with a higher proficiency level are likely to have higher JSST scores. The JLPT scores are not correlated with ESST scores (Multiple r=0.39). This means higher L1 WMC does not necessarily promote a higher proficiency level in Japanese. This could be because the tasks in the Japanese proficiency tests and ESST are very different.

Table 1. Summary of the test scores

Learner #	Level	ESST(100)	JSST(100)	PT(50)	MJT(50)	JLPT(100)
1	5	90	69.3	50	48.8	98.8
2	5	85	67.1	42.5	50	922.5
3	4	78.3	55	22.5	46.4	68.9
4	5	71.7	54.3	43.8	48.8	92.6
5	5	66.7	64.3	43.8	46.4	90.2
6	4	65.8	47.1	36.3	46.4	82.7
7	5	65	49.3	37.5	42.9	80.4
8	5	63.3	55	41.3	46.4	87.7
9	5	60	42.9	43.8	48.8	92.6
10	3	60	37.1	22.5	35.7	58.2
Average		70.6	54.1	38.4	46.1	84.4

The average number of produced sentences in JSST was 45.1 sentences per learner; of those, the correct sentences were 40.7 (90.2%). Most of their utterances were mono-clausal (91.6%) and null subjects were used 71.3% of the time. Interrogative forms were sometimes produced (14.7%). One of the most common errors was particle misuse (1).

(1) a. Keitai de takusan apuri ga arimasu. (de → ni)
cellphone LOC many apps NOM there is
'There are many apps on my cellphone.'

b. Honya ni hon o kai mashoo. (ni → de)
bookstore LOC books ACC buy let's
'Let's buy a book at a bookstore.'

c. Jitensha o noru no ga suki-desu (o → ni)
bicycle ACC to ride DAT NOM like
'I like riding on a bicycle.'

Other errors include lexical choices and some usage errors (2).

(2) a. Kaigi ni hairu.
Meeting LOC enter
'I attend a meeting.'

b. Ichiban osusume-na omise wa nandesu ka.
most recommended store NOM what Q
'Where is the shop you recommend the most?'

In sentence (2a), *hairu* 'to enter' should be either *iku* 'to go' or *deru* 'to attend.' In sentence (2b), *osusume-na* should be *osusume-no,* and *nan* 'what' should be either *doko* 'where' or *donna omise* 'what kind of shop.' Particle drops are often observed in Japanese speech. Therefore, sentences like (3) were not considered ungrammatical.

(3) a. Kono kooen shizuka desu-ne.
this park quiet
'This park is quiet, isn't it?'
b. Kono kanji doo yomimasu ka.
This kanji how read Q
'How do you read this kanji?'

Prosodic errors do not affect JSST scores. However, the difficulty of prosodic features are reported (Aysawa 2000; Goss and Nakayama 2011). Thus, the production data were also analyzed to see prosodic features, such as mispronunciation and downsteps. Although there was not much, some mispronunciations, especially for English loanwords, were observed (4).

(4) a. Nyuuzu wa amari mimasen. (nyuuzu → nyuusu)
news NOM much not watch
'I don't watch the news much.'
b. *Benrina appuri o tsukau.* (appuri → apuri)
convenient apps ACC use
'I use convenient apps'

Venditti (2006) claims that understanding prosody is an important element in speech. Downstep is one of the most important prosodic features in the Japanese language. It is a phonological process in which F0 declines gradually after the first highest pitch (Venditti 2006). Thus, it would be beneficial to include a prosodic feature like downstep as an important component for fluency and examine it in L2 Japanese learners' speech production. Most of the learners only successfully produced downsteps less than 50% of the time (Table 2). This confirms the difficulty of acquiring the prosodic feature in L2 Japanese even among intermediate and advanced learners. However, note that the regression analysis did not show a correlation between the JSST scores and the number of downsteps (Multiple r= 0.25).

Table 2. Downstep

Learner #	Correct sentences	Downstep	Downstep (%)
1	53	24	45.2%
2	48	21	43.8%
3	48	17	44.7%
4	41	33	80.5%
5	45	10	22.2%
6	36	21	58.3%
7	42	17	40.5%
8	43	13	30.2%
9	29	11	37.9%
10	32	17	53.1%
Average	40.7	18.4	45.2%

6. Conclusion

The relationship between WM and individual language skills has received much more attention recently. Though preliminary, the current study presented a newly created L2JSST and documented its correlations with ESST scores and Japanese proficiency test scores. Individuals with larger WMC in English tend to have larger WMC in Japanese. These findings appear to indicate that the Japanese version of SST that we created is very promising and functions as a viable test. However, despite the positive outcomes of the present study, there are limitations to be noted. It may be too early to conclude the validity of JSST and its correlation with WMC due to the small number of participants. This is something to address in the future if findings from this study are to be generalized. Utilizing the JSST in future trials with additional participants can help shed more light on the validity of the test and the relationship between WMC and Japanese proficiency level. Finally, the learners of this study produced grammatical sentences with a relatively high correction rate with the JSST. When prosody is considered, however, proper downstep was observed in less than half of the correctly produced sentences. The acquisition of prosody seems challenging for speakers of intermediate to advanced Japanese.

References

Ayusawa, Takako. (2003) Acquisition of Japanese Accent and Intonation by Foreign Learners. *Journal of the Phonetic Society of Japan* 7.2, 47–58.

Daneman, Meredyth and Ian Green. (1986) Individual Differences in Comprehending and Producing Words in Context. *Journal of Memory and Language* 25.1, 450-466.

Goss, Seth, and Mineharu Nakayama. (2011) Prosody and Comprehension in Oral Reading by L2 Japanese Learners. *Second Language* 10, 33-50.

Fortkamp, Malice Borges Mota. (1999) Working Memory Capacity and Aspects of L2 Speech Production. *Communication and Cognition* 32, 259-296.

Fortkamp, Malice Borges Mota. (2000) Working Memory Capacity and L2 Speech Production: An Exploratory Study. Unpublished Ph.D. dissertation, Federal University of Santa Catarina, Florianópolis, Santa Catarina, Brazil.

Itomitsu, Masayuki and Mineharu Nakayama. (2005) Japanese Language Proficiency Test: Practice Test. The Ohio State University. Memory Capacity, Noticing, and L2 Speech Production. *Singo* 32, 52.

Lee, Albert and Peggy Mok. (2018) Acquisition of Japanese Quantity Contrasts by L1 Cantonese Speakers. *Second Language Research* 34.4, 419-448.

Maki, Hideki, Jessica Dunton and Carlyn Obringer. (1999) What Grade Would I Be in If I Were Japanese? Paper Presented at *The 14th Annual Conference of the Association of Teachers of Japanese*.

Mota, Malice Borges. (2003) Working Memory: Capacity and Fluency, Accuracy, Complexity, and Lexical Density in L2 Speech Production. *Fragmentos, Número* 24, 64-104.

Venditti, Jennifer. (2006) Prosody in Sentence Processing. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *Handbook of East Asian Psycholinguistics Vol.2: Japanese*. 208-217. Cambridge, UK: Cambridge University Press.

Weissheimer, Janiana and Malice Borges Mota. (2009) Individual Differences in Working

Memory Capacity and the Development of L2 Speech Production. *Issues in Applied Linguistics* 17.2, 93-112.

hattori.12@osu.edu

Appendix

Japanese speaking span tests (sample words)
ホテル (*hoteru*: hotel), 薬 (*kusuri*: medicine), 公園 (*kooen*: park), アパート(*apaato*: apartment), 文化 (*bunka*: culture), 飲み物 (*nomimono*: drink), 病院 (*byooin*: hospital), スマホ (*sumaho*: smartphone), 歴史 (*rekishi*: history), 正月 (*shoogatsu*: New Year), 復習 (*hukushuu*: review), 鞄 (*kaban*: bag), 銀行 (*ginkoo*: bank)