The Effect of Oral Repetition on L2 Speech Fluency:
An Experimental Tool and Language Tutor

Yuki Yoshimura, Brian MacWhinney

Language Technologies Institute & Department of Psychology
Carnegie Mellon University
Pittsburgh, Pennsylvania U.S.A.
`yyuki@cmu.edu, macw@cmu.edu`

Abstract

This paper discusses the effects of oral repetition and practice in improving speech fluency for adult second language acquisition. The experiment examined the impact of practice on fluency in learning Japanese as a second language. The measures included read-aloud time and speech production time during sentence rehearsal. The results indicated gradual improvement of speech fluency as the number of practice trials increases. Implications and directions for technology to promote fluency are discussed from the perspective of psycholinguistic research.

Index Terms: speech fluency, repetition, second language learning, Japanese.

1. Introduction

Oral repetition and imitation-based practice has been widely used as one of the major methods to improve speech fluency in second language (L2) learning. Such rehearsal is used to enhance the familiarity of novel words, phrases, and sentences with an emphasis on intonation and speed. The use of oral repetition has proven to be useful and can be supported theoretically as well. Studies of individual differences in language learning have shown that it is important to maintain and rehearse phonological information in working memory [1]. The linking of phonological short-term memory with long-term memory in language learning is crucial to triggering the chunking of lexical and syntactic units to promote fluency [2-4].

It is generally believed that learners have more problems with sentence production when a sentence contains novel words or unfamiliar phrases. Considering the fact that novel information increases cognitive loads in working memory, language fluency is more likely to be interrupted when L2 learners have to process new words. In the foreign language classroom, teachers often ask learners to repeat sentences. However, in the classroom, it is difficult for a teacher to pay attention to individual students to point out errors in their speech. Some learners are able to repeat a model sentence at the same speed immediately, while others may pretend that they are done with the repetition when the classroom becomes silent. Individual differences may vary even more with increases in the number of novel vocabulary items in a sentence. Practice can reduce these disfluencies. However, no study has looked at the possible relation of simple oral repetition with fluency in L2 speech. The current study looked at the effect of oral practice in rehearsing sentences that involved novel vocabulary.

2. Automaticity and L2 Speech Fluency

There is a strong link between automatic processing and fluency [5, 6]. Researchers in the field of skill acquisition [7, 8] hold that, with practice, a skill moves from controlled to automatic processing. Research has also indicated that higher working-memory capacity brings advantages in fluency [9-11]. Regardless of individual working-memory capacity, the automatic processing of language frees up capacity for other information, which results in fluent use of language. At the same time, higher working-memory capacity allows speakers to process larger pieces of information fluently. Automatic processing can be contrasted with controlled processing. Schneider & Shiffrin (1977) and Shiffrin & Schneider (1977) characterized an automatic process as activating a sequence of nodes uniformly and obligatorily for a certain input. Activation of this sequence requires no conscious attention or active involvement. On the other hand, controlled processing involves the activation of nodes that are not always active for a certain input. As a result, activation of a particular set of nodes requires conscious attention. A core goal in L2 learning is to enhance the automaticity of linguistic information so that learners can handle as much L2 material as possible in a fluent fashion. Automatization in L2 learning describes the phrase of information speed-up or representational change of knowledge or sometimes both [12].

Studies of language production, such as MacKay (1982) [13], consider speed-up as a core factor for automaticity [14]. When automaticity is viewed as a part of the development of oral language fluency, speed-up plays an important role. Unlike language comprehension, oral production involves highly proceduralized motor skill components. For these articulatory skills, speed-up is thought to depend on processes of chunking and automatization. In interactive activation models speed-up is characterized in terms of strong and direct activation between units [15, 16].

The current experiment examines how changes in automaticity based on lexical novelty can influence fluency.

3.   Method

1.1. Participants

Participants were 30 learners learning Japanese as a foreign language at Carnegie Mellon University. All participants were enrolled in the course of Intermediate Japanese and had studied Japanese for the total of 10 months.

1.2. Stimuli

Each sentence consisted of 6 content words, along with attendant function words. The syllable length for each sentence varied between 25 and 31, and the target novel words consisted of 2, 3 or 4 syllables. The design involved the factor of novel words with four levels (zero, 1, 2 and 3), repetitions (from 1 to 6), and a task factor that contrasted the six repetition trials with a final sentence retrieval task. The basic structure was an affirmative sentence using a canonical word order (SOV or OV) in Japanese with a verb and an object that comes with some adjectives, and places.

Example sentence stimuli:
chichi wa denwa de keizai no *tokucho* o hanashi-mashita.
[chichi=father, wa=topic/subject marker, denwa=phone, de=by, keizai=economics, no=of, tokucho=characteristics, o=object marker, hanashi-mashita=talked].

Each sentence was designed to make sense by itself without contexts. The words were selected from the course textbook and were selected to be familiar both in oral and written forms. The novel words as target items were selected from words that never appeared in the textbooks or materials used in the classrooms. To make a clear distinction between familiar words and novel words, the familiar words were all taken from the Elementary level. Novel words were inserted in the middle of each sentence to avoid primacy and recency effects.

Words commonly used in hiragana (phonetic-based character) were displayed in hiragana. Words commonly used in kanji (Chinese character) were displayed in kanji with a phonetic guide in hiragana so that participants would not have any problems with reading. In addition, novel words were underlined with an English translation below. Target items were either nouns or adjectives.

1.3. Procedure

The experiment consisted of a sequence of three tasks; listening, reading aloud, and sentence-retrieval production from memory. Each sentence appeared on a computer screen. The participants had a chance to listen to each sentence as a model speech for three times repeatedly at the beginning of the experiment. After they heard the model sentence, participants were asked to read it aloud six times. The numbers of 1, 2, 3, 4, 5 and 6 appeared on the top of the screen to indicate how many times participants had read aloud each sentence. During reading aloud, participants were instructed to memorize the sentence to be able to repeat it later. After six oral repetitions, the sentence stimuli disappeared from a computer screen, and participants were asked to retrieve the sentence with a delay of 1000 ms. Participants had a chance to go through the practice session of the experiment to understand the whole procedure before they started the actual trial.

This task has a couple of key features. By providing written information, the read-aloud task eliminated problems involved with the participants' limited listening ability. Also, by practicing each sentence orally, the phonological information for both familiar and novel items could be activated in fluent rhythmic speech.

To avoid disfluencies, it is important to reduce working-memory load when to-be-processed items involve novel information that requires controlled processing. Processing a novel word consumes a lot of cognitive resources that may cause disfluency by affecting the rest of the familiar words in a sentence production. To eliminate other working-memory loads such as lexical and conceptual decisions [17], a task that requires spontaneous phase of speech was not used. Instead, this procedure focuses on the automatization of phrasal combinations as major method for improving fluency.

4.   Results

The length of utterance in terms of read-aloud time and sentence-retrieval time was measured from the beginning of utterance to the end of utterance. The read-aloud time as oral practice was measured for all 6 cycles. The analysis was carried out using digitally recorded audio in Cool Edit 2000.

Analysis of variance showed significant main effect for the factor of novel words by longer utterances for sentences with larger number of novel words, $F(3, 87) = 23.107$, $p < .001$. The measure of 6 cycles as oral practice was also significant where time decrease in length of utterances was observed; $F(5, 145) = 32.243$, $p < .001$. Figure 1 illustrates the clear decline in read-aloud time from cycle 1 to cycle 6. Further detailed analysis by pairwise comparisons showed the steepness of time reduction through repetition. The time difference from oral practice from cycle 1 to cycle 2, from cycle 3 to cycle 4, and from cycle 5 to cycle 6 were not significant, but the reduction from cycle 2 to cycle 3, and from cycle 4 to cycle 5 were both significant (see Table 1).

Table 1: Mean difference of read-aloud time among 6-cycle oral practice.

Cycle number	Mean Difference	Std. Error
1-2	.164	.111
2-3	.288 *	.076
3-4	.116	.082
4-5	.385 *	.076
5-6	- .075	.069

*$p < .01$
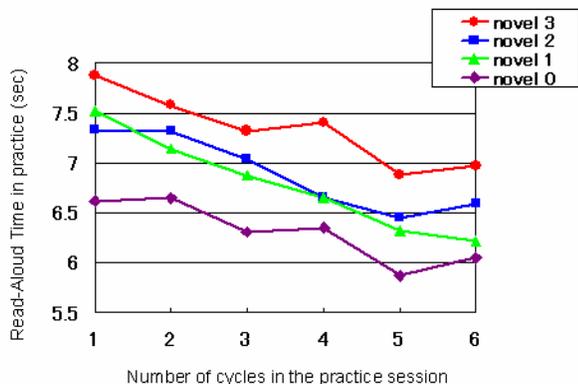
Figure 1: Learners' performance on the length of read-aloud time in the oral practice session.

Figure 2: Learners' performance on the average length of read-aloud time in the oral practice session and production time from memory retrieval.

Across the 6 cycles of reading aloud, sentences with three novel words initially took longer than sentences with zero novel words. To quantify the learning effect, a trend analysis was conducted comparing the linear component across trials for each condition. The analysis of within-subjects orthogonal polynomials showed that the linear component of time reduction from cycle 1 to cycle 6 was largest for sentences with a single novel word. The linear reduction of the condition of novel-1 sentences vs. other conditions was significant at $F(1, 29) = 6.628$, $p = .05$. The condition of novel-1 differed from novel-zero condition ($F(1, 29) = 8.104$, $p = .05$), and also from the mean of novel-2 and novel-3 ($F(1, 29) = 4.265$, $p = .05$). The linearity of novel-zero condition also illustrates a gentle slope that differed from the rest ($F(1, 29) = 4.393$, $p = .05$). The rest of the comparisons showed no difference. The result can be understood as resulting from two forces. First, sentences with no novel words do not undergo much speed-up, because they are already relatively easy. Second, sentences with two and three novel words do not undergo much speed-up, because they are too difficult to automatize fully with only six repetitions.

A two-way analysis of variance using the factors of novelty (0, 1, 2 and 3) and task (read-aloud and sentence-retrieval) was used to analyze the difference between read-aloud time and sentence-retrieval time after 6 cycles of oral practice. The analysis produced a significant main effect for novelty $F(3, 87) = 7.388$, $p < .01$, which suggests that sentences with more novel words require longer processing time. Also, read-aloud time was significantly shorter than sentence-retrieval time; $F(1, 29) = 40.360$, $p < .01$. These patterns of time difference between read-aloud and sentence-retrieval task is described in Figure 2. To further determine the difference between read-aloud and sentence-retrieval time in detail, a paired samples t-test was performed where the analysis compares the mean of each individual participant's length of utterances between read-aloud and sentence-retrieval time. The analysis illustrated significantly shorter length of utterances for read-aloud time than for sentence-retrieval time for three conditions of 1, 2 and 3 of novel words respectively at $t(29) = -4.908$, $p < .0025$, $t(29) = -4.995$, $p < .0025$, $t(29) = -4.969$, $p < .0025$ (alpha corrected by Bonferroni correction). The zero novel word condition also showed a significant difference at $t(29) = -3.076$, $p < .005$.
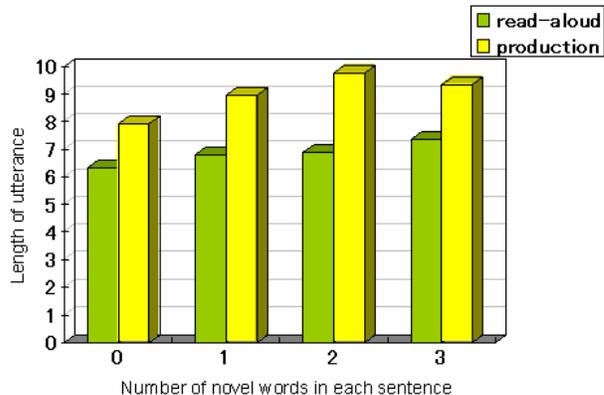
5. Discussion

The results obtained in this study indicate that oral practice significantly increases fluency. The fact that the read-aloud time decreased significantly over six cycles showed the effect of oral practice for speech production when reading aloud. This speed-up in language performance is predicted by the theory of automaticity in skill acquisition. There was a less time reduction with three novel words than with one novel word. This pattern indicates that more practice is required to enhance automaticity for larger pieces of novel information. Although consistent time reduction was observed for all four conditions, none of the conditions reached full automaticity. Even sentences with one novel word could have benefitted from additional practice and repetition. From the results of the current study, we cannot determine exactly how much practice is needed to achieve full fluency.

In addition to reading aloud, the experiment involved listening to a model speech for three times as input. Participants in the task heard a sample speech before they start rehearsing each sentence orally. If only receiving aural input in terms of listening to speech samples was sufficient to enhance speech fluency, we would not see any time reduction during the task of oral rehearsal from cycle 1 to 6, since they should be able to perform well from cycle 1. This supports the cliam that oral repetition improves fluency. However, this study does not provide an independent estimate of the amount of this speed up that can be attributed to having heard each sentence modeled three times. It would be useful to have additional experiments that allow analyses to compare the effect of aural input and oral practice separately in the near future. First, it would be valuable in language education, if we can determine the most appropriate frequency of aural input and oral practice and also the best combination of both. Second, there are various ways to control the order of input and practice. The task in the current study used 6 repetitions of oral practice followed by three repetitions of aural input, but it is possible that alternating aural input and oral practice by turns would be better than having these procedures applied in blocks. Finally, it is important to test whether the immediate effects of this training are maintained in the long run.

6. Directions for Technology Use

It is difficult in the framework of a lab study of this type to provide practice and training on a daily basis. However, language learning is an outcome of continuous daily practice. It is important to assess how these methods function in the classroom context. To do this, we need a computer-based system that works both as an experimental tool and as a learning tutor that can be adjusted for different levels of individual learners.

An ideal system should be able to provide multiple practice methods that learners can choose. These multiple methods work as an experimental conditions at the same time. For example, it is useful to have a tutor which has four methods that involve different sequence or order of input and oral practice. Method A provides listening to a sentence as input and oral practice alternatively for one sentence, whereas method B provides input only at the beginning repeatedly that follows repetitive oral practice. In methods A and B, learners practice only each sentence repeatedly until they achieve a certain level of fluency. In other words, the next sentence does not appear until learners show sufficient fluency. Methods C and D are similar to methods A and B in ordering aural input and oral practice, but differ in ordering sentences. Learners practice multiple sentences in Methods C and D simultaneously where fluency level of all sentences are supposed to increase at the same time. Detailed conditions are illustrated in Figure 3.
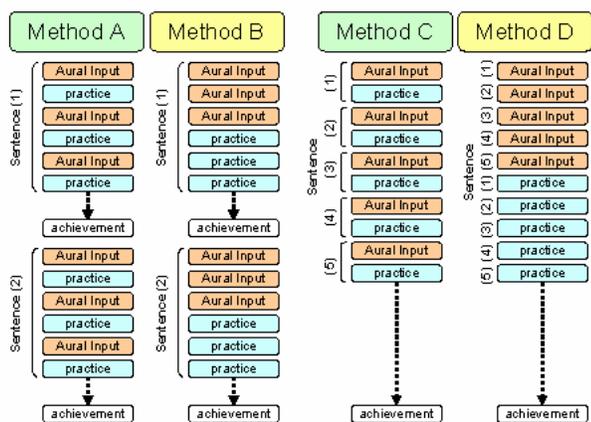

Figure 3: An ideal system works both as an experimental tool and a tutor.

For a system to work both as an experimental tool and as a language tutor, the system must be able to handle three things automatically. First, it must be able to collect and record learners' language production as computerized data files. Second, it should be able to recognize learners' speech during the session of practice and estimate how much it is similar to a model speech as an achievement level. If this cannot be done accurately, it may be necessary to have learners help in making these judgments. Finally, based on the system's evaluation of the learners' progress, the number of repetition trials in accord with their estimates should be adjusted. The development of such a system can make a major contribution to L2 education by helping students improve oral fluency.

7. References

[1] Gathercole, V. and A. Baddeley, *Working memory and language*. 1993, Hillsdale, NJ: Lawrence Erlbaum Associates.

[2] Papagno, C., T. Valentine, and A. Baddeley, *Phonological short-term memory and foreign-language vocabulary learning*. Journal of Memory and Language, 1991. 30: p. 331-347.

[3] Ellis, N., *Sequencing in SLA: Phonological memory, chunking, and points of order*. Studies in Second Language Acquisition, 1996. 18: p. 91-126.

[4] Gupta, P. and B. MacWhinney, *Vocabulary acquisition and verbal short-term memory: Computational and neural bases*. Brain and Language, 1997. 59: p. 267-333.

[5] Schneider, W. and R. Shiffrin, *Automatic and controlled information processing in vision*, in *Basic processes in reading: Perception and comprehension*, D. Laberge and S. Samuels, Editors. 1977, Lawrence Erlbaum Association: Hillsdale, N. J.

[6] Shiffrin, R.M. and W. Schneider, *Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory*. Psychological Review, 1977. 84: p. 127-190.

[7] Logan, G.D., *Toward an instance theory of automatization*. Psychological Review, 1988. 95 (4): p. 492-527.

[8] Newell, A. and P.S. Rosenbloom, *Mechanisms of skill acquisition and the law of practice.*, in *Cognitive skills and their acquisition*, J.R. Anderson, Editor. 1981, Hillsdale: NJ: Erlbaum.

[9] Gathercole, S.E. and A.D. Baddeley, *Working memory and language*. Essays in cognitive psychology. 1993, Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc. xiii, 266.

[10] Just, M.A. and P.A. Carpenter, *A capacity theory of comprehension: Individual differences in working memory*. Psychological Review, 1992. 99: p. 122-49.

[11] Osaka, M. and N. Osaka, *Language-independent working memory as measured by Japanese and English reading span tests*. Bulletin of the Psychonomic Society, 1992. 30: p. 287-289.

[12] Segalowitz, N. and J. Hulstijn, *Automaticity in Bilingualism and Second Language Learning*, in *Handbook of bilingualism: Psycholinguistic approaches*, J.F. Kroll and A.M.B. de Groot, Editors. 2005, Oxford University Press: New York, NY, US.

[13] MacKay, D.G., *The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior*. Psychological Review, 1982. 89: p. 483-506.

[14] DeKeyser, R., *Automaticity and automatization*, in *Cognition and Second Language Instruction*, P. Robinson, Editor. 2001, Cambridge University Press: Cambridge. p. 125-151.

[15] Dell, G., *Speaking and misspeaking*, in *An invitation to Cognitive Science: Language*, L. Gleitman and M. Liberman, Editors. 1995, MA: MIT Press.: Cambridge. p. 183-208.

[16] Stemberger, J., *An interactive activation model of language production*, in *Progress in the psychology of language*, A.W. Ellis., Editor. 1985, Erlbaum: Hillsdale, NJ.

[17] Levelt, W.J.M., *Speaking: From intention to articulation*. 1989, Cambridge, MA: MIT Press.