Reading in a Foreign Language October 2023, Volume 35, No. 2
ISSN 1539-0578 pp. 270–292

The Effects of Chunk Reading Training on the Syntactic Processing Skills and Reading Spans of Japanese Learners of English

Takumi Kosaka
The University of Melbourne
Australia

Abstract

The present study investigated the impact of chunk reading training (CRT) on the online syntactic processing and verbal working memory (WM) of Japanese EFL (English as a foreign language) learners in secondary school. For four weeks, the treatment group (N = 31) underwent CRT, while the control group (N = 25) participated in reading training in block format. A reading span test (RST) was administered as the pretest, posttest, and delayed posttest in this study to assess learners' online syntactic processing and verbal WM. The results showed that syntactic processing and verbal WM increased only in the treatment group after training, but the differences between the two groups were not statistically significant. Nevertheless, these results suggest that CRT has the potential to positively affect the development of learners' online syntactic processing skills and verbal WM.

Keywords: Chunk-and-Pass Model, usage-based approaches, construction grammar, chunk reading training, reading span test, verbal working memory, online syntactic processing, Japanese EFL learners, secondary school EFL learners

Drawing from usage-based approaches (e.g., Tomasello, 2003) and construction grammar (e.g., Goldberg, 2019), Christiansen and Chater (2016) developed the Chunk-and-Pass Model, a novel psycholinguistic model for language processing and acquisition. The Chunk-and-Pass Model posits that because humans have memory limitations ranging from 7 ± 2 (Miller, 1956) to 4 ± 1 (Cowan, 2001) pieces of information, language users must recode language inputs into larger, more abstract representations so that they do not interfere with preexisting information (Christiansen & Chater, 2016). For instance, the number of words in the sentence *I went to the grocery store yesterday and bought some vegetables for dinner* (13 words) exceeds the memory limitation of 7 ± 2 (Miller, 1956), which may cause the earlier information in the sentence to be lost. To cope with the challenges of rapid and transient language input, first language (L1) speakers of English utilize a process of chunking, whereby individual words are chunked together and passed up to a higher level of representation (i.e., from the word level to the multiword level) to create an abstract but singular whole.

Another key assumption of the Model is that language acquisition is akin to skill learning. This implies that learning how to process language inputs via chunk-and-pass processing requires continuous practice (Christiansen & Chater, 2016). This process requires the learner to have schematic knowledge of linguistic constructions (Goldberg, 2019). The proposal of the Chunk-and-Pass Model caused widespread reactions in the language sciences (e.g., Clark, 2018; Guo & Ellis, 2021). The present research was conducted as an initial step toward the Chunk-and-Pass Model's application in the instructed learning of English as a foreign language (EFL).

Literature Review

The Chunk-and-Pass Model

Usage-based approaches (e.g., Tomasello, 2003), especially construction grammar (e.g., Goldberg, 2019), posit that language acquisition involves attaining knowledge of linguistic constructions that map language forms and meanings through language experience. In English, the simple "subject (S) + verb (V) + object (O)" syntactic structure conveys the abstract meaning "who did what to whom" (Lupyan & Christiansen, 2002). Further, the ditransitive construction "V + O1 + O2" (e.g., *give me a present*) presents the schematic meaning "X causes Y to receive Z" (Goldberg, 2019). Construction grammar claims that every linguistic construction has its own schematic meaning(s) (Goldberg, 2019) and that language learners extract the schematic representations of linguistic constructions by mapping their forms to their functions through language experience (Diessel, 2015).

Building on usage-based approaches and construction grammar, Christiansen and Chater (2016) developed a novel psycholinguistic model called the Chunk-and-Pass Model to explain language processing and acquisition. According to the Model, since the memory limitations of humans range from 7 ± 2 (Miller, 1956) to 4 ± 1 (Cowan, 2001) items, language users must chunk language inputs and recode them into larger units of linguistic sequences that form more abstract levels of linguistic representations; otherwise, information obtained earlier will be overridden by information obtained later (Christiansen & Chater, 2016). Miller (1956) hypothesized that there is a maximum memory capacity of nine items; however, a sentence in English (and other languages) typically contains more than 10 words. Thus, language inputs are subject to a fundamental memory limitation that is commonly referred to as the now-or-never bottleneck (Christiansen & Chater, 2016); that is, if a language user is unable to process individual words at a given moment, the earlier words are lost and cannot be recovered. To cope with rapid and transient language inputs, language users must recode language information into larger, more abstract representations (e.g., going from the word level to the multiword level) via chunk-and-pass processing (Christiansen & Chater, 2016).

An outstanding hypothesis of the Model is that language acquisition is akin to skill learning. It involves learning how to appropriately chunk language inputs through continuous practice (Christiansen & Chater, 2016). For efficient chunk-and-pass processing, language users must have access to schematic knowledge of linguistic constructions, as these constructions serve as computational procedures for creating chunks; thus, language learners must first develop schematic knowledge (Christiansen & Chater, 2016). However, because the now-or-never bottleneck implies that language information must be processed and stored within a limited time frame, it is essential that most of the learning takes place during this immediate processing period; otherwise, the processed and stored representations will be lost forever. That is, learning takes place in real time; this is referred to as "*online* learning" (Christiansen & Chater, 2016). In short, language learners need to learn how to chunk their respective target languages based on their acquisition of schematic constructions, and the required acquisition and processing skills can be gained through continuous *online* (real time) practice.

Chunk Reading Training

As discussed in the previous section, recent psycholinguistic research has highlighted the importance of chunking for language processing and acquisition, particularly in light of the principles of construction grammar. In the last two decades or so, educational researchers have used the approach of chunking sentences to improve students' reading comprehension (i.e., chunk reading training [CRT]). For example, Walker et al (2005) developed a visual-syntactic text formatting (VSTF) software for English sentence segmentation, specifically to facilitate the formation of sentences within one or two fixation eye spans (8–30 characters), thereby increasing learners' syntactic awareness (Park et al, 2019; Park & Warschauer, 2016). Walker et al (2005) used VSTF with students of English as a second language (ESL) in a secondary school in the United States and analyzed their test scores for the reading section of a district standardized test (Terra Nova). The control group received normal reading practice in the traditional block format, in which sentences were presented in paragraphs rather than broken into smaller multiword units. The results showed that the ESL students in the VSTF group significantly improved their reading skills, whereas the students in the control group did not. More recently, Park et al (2019) reported that VSTF effectively improved the syntactic awareness of fourth- and sixth-grade L1-English and ESL students in California.

Although there is a growing body of evidence for the effects of VSTF on reading comprehension, i.e., offline processing (McMaster et al, 2015), little has been reported on whether CRT develops one's online processing skills. Furthermore, the chunking algorithm used in the VSTF software has syntactic and semantic inconsistencies when viewed from the perspective of construction grammar and the Chunk-and-Pass model. For example, the VSFT software would segment the sentences below as follows (Park et al, 2019, p. 7):

(1) *There, / she earned a degree / in physics*.
(2) *After college, / she had / a decision / to make*.

In sentence (1), the subject and the verb phrase are chunked together (*she earned a degree*), whereas in sentence (2), the subject and the verb are separated from the object phrase (*she had / a decision*). In a study by Jolsvai et al (2020, p. 6), L1-English speakers had slower reaction times to multiword units that were both syntactically and semantically fragmented (e.g., *without the primary*) than to phrasal items (*nothing to wear*), as fragmental trigrams tended to be more difficult for them to process through Chunk-and-pass Processing (i.e., by creating a singular representation). This means that, from the perspective of the Chunk-and-Pass model, learners would find it difficult to process grammatically and semantically fragmental multiword units.

Challenges specific to Japanese EFL learners/readers. In addition, the separation of the verb and object in a sentence may make it challenging for Japanese ESL/EFL learners/readers to develop their online sentence processing skills. Fender (2003) reported that, compared to Arabic head-initial verb phrases (VSO/SVO) learners of English, Japanese (SOV) learners of English would find it considerably difficult to integrate a verb and object into a phrase. Fender (2003) postulated that Japanese speakers would have acquired word integration skills that reflect the head-final phrasal structure of their L1 (SOV word order), leading to difficulties when they attempt to integrate a verb and an object in English. It should be noted that the reading comprehension levels of the Japanese and Arabic ESL learners included in Fender's (2003) study were not statistically different. Consequently, reading training that separates verbs and objects may make it difficult for Japanese ESL/EFL learners of English to acquire the ability to integrate a verb and an object into the SVO word order.

As mentioned above, the Chunk-and-Pass Model posits that language acquisition involves learning how to process language inputs through a continuous practice of language processing (Christiansen & Chater, 2016) which would be offered in the activity of reading. The presentation of inconsistent and fragmented multiword units may hinder second language (L2) acquisition due to the nature of language processing and acquisition, as given in the Chunk-and-Pass Model.

The following section explores the use of the reading span test (RST) as a measure of learners' online syntactic processing skills and verbal working memory, which was applied in the present study to address the research questions.

Working Memory and the Reading Span Test

This section outlines the use of the RST, a working memory test, as a measure of language processing ability. Working memory (WM) is a fundamental cognitive process responsible for the simultaneous processing, storage, and retrieval of information (Baddeley & Hitch, 1974). There is mounting evidence that WM plays an important role in language processing (e.g., Gao & Wu, 2023), comprehension (e.g., Peng et al, 2018), and language acquisition (e.g., Perez, 2020).

The RST (Daneman & Carpenter, 1980) is the most widely used measure of verbal WM. This test requires participants to simultaneously process and store information by reading unrelated sentences and memorizing the last words in the sentences. The test sessions and sentence stimuli vary in number from study to study, but in general, sentence stimuli are presented in 2–6 sentence conditions, and each condition involves 3–5 WM sessions. After reading a certain number of sentences, participants are prompted to recall as many words as possible. In an RST for L2 learners, a grammaticality judgment task (GJT) (Harrington & Sawyer, 1992) is generally included as a measure of participants' syntactic processing capacity, while the word recall task (WRT) is employed to evaluate their capacity for storing and retrieving information. Grey et al (2015, p. 144) conducted a study in which a Spanish version of the GJT was administered to college-level native English speakers learning Spanish to measure their implicit syntactic knowledge before and after a five-week intensive language experience abroad. The task included grammatical and ungrammatical sentences, and the word order for sentences in the ungrammatical condition was randomized as follows:

(3) Grammatical: *Tengo que [correr muchas millas] esta semana.*
(4) Ungrammatical: *Tengo que [millas muchas correr] esta semana*.

The authors reported that the accuracy rate of the participants increased significantly from 49.03% to 61.92% after they participated in the study abroad program.

Theoretically, it has been argued that the RST can measure individual variations in WM, assuming that WM is functionally independent of the learner's linguistic knowledge (e.g., Just & Carpenter, 1992). However, the RST has been criticized for its reliance on language experience and proficiency (MacDonald & Christiansen, 2002). In this regard, it has been suggested that participants with high levels of language experience may have high WM capacities due to their superior language processing abilities (MacDonald & Christiansen, 2002). Farmer et al (2017) demonstrated this hypothesis in their empirical study. They recruited L1 speakers of English to take an automated RST as well as the author recognition task, which measures linguistic experience. The results revealed a significant correlation between reading experience and reading span: Individuals with relatively high levels of reading experience had high reading spans, which could be attributed to their language processing abilities.

Notably, individual reading spans depend on language proficiency when the RST is administered to L2 learners. Van den Noort et al (2006) recruited L1 speakers of Dutch who had learned German as an L2 and Norwegian as a third language (L3) and examined their reading spans for each language. The results showed that the participants' reading spans were highest for their L1, followed by their L2 and L3 (i.e., L1 > L2 > L3), indicating that WM, as measured by the RST, is influenced by the learner's language proficiency.

Based on the Chunk-and-Pass Model, Christiansen and Chater (2016) argued that "the reading span test is simply another measure of language processing skills" (p. 187).[1] It is hypothesized that learners with superior language processing abilities are able to achieve higher word recall rates than learners with weaker language processing abilities because they are able to generate abstract representations through chunk-and-pass processing, thus allowing more cognitive resources to be devoted to lower levels of representation (e.g., the word level in the context of the RST). Based on the assumption underlying the Chunk-and-Pass Model, an RST including a GJT were employed in the present study to investigate the effects of CRT on Japanese EFL secondary school learners' online syntactic processing skills and verbal WM.

The Present Study

As discussed in the previous sections, the Chunk-and-Pass model emphasizes the importance of chunking for efficient language processing and acquisition. Although educational research has leveraged chunking for reading comprehension, early studies on CRT often lacked a solid linguistic or psycholinguistic framework to support their decisions on how chunks are formed

[1] However, they also caution that language processing skills are supported by many other factors including a host of domain-general components, and thus, there should not be a one-to-one mapping between reading spans and language processing skills/experience (p. 187).

from different types of linguistic constructions. No study has adopted the Chunk-and-Pass model (Christiansen & Chater, 2016) to provide a rationale for the use of chunk-based training to support language processing and acquisition. It would be perplexing to learners if they were exposed to an inconsistent chunking system of sentence segmentation. Furthermore, existing studies on CRT have only measured training effects using regional/state standardized tests of reading comprehension (e.g., Park et al, 2019). No study has provided a more direct measure of whether CRT promotes learners' implicit and online syntactic processing abilities and whether CRT helps to expand learners' verbal WM. As we know from previous psycholinguistic studies (e.g., Fender, 2003), L2/EFL learners' reading comprehension skills are not of the same construct as their online sentence processing abilities. Therefore, the present study aimed to address the following research questions:

RQ1: Does CRT help develop the implicit online syntactic processing skills of Japanese EFL learners?

RQ2: Does CRT help expand the verbal WM of Japanese EFL learners?

The aim of the present study was not to investigate the superiority of one type of text formatting, as in previous studies on VSTF, but to elucidate the effects of CRT *per se* on the syntactic processing skills and verbal WM of Japanese EFL learners. Therefore, although the study followed a within- and between-subjects design involving a treatment group and a control group, it should be noted that the CRT approach used was manipulated by the author, and some of the manipulations were not considered in the development of a training program for the control group.

Method

Participants

Initially, 69 second-year students (16–17 years old) from a Japanese secondary school were recruited and divided into treatment and control groups. They had studied English as a foreign language in Japan since age 13, had no previous experience of living in English-speaking nations, and spoke Japanese as their L1. Their English proficiency levels were low (levels A1–A2) according to the *Common European Framework of Reference for Languages* (Council of Europe, 2001). Thirteen students were subsequently excluded because they did not complete over 70% of the training sessions, resulting in 56 participants (treatment: 31, control: 25). Due to COVID-19-related constraints, additional proficiency tests could not be administered. Instead, the latest midterm exam, which encompassed the areas of grammar and reading comprehension, was used to gauge the students' English proficiency levels, in line with Zhang and Dong's approach (2019). No significant difference in proficiency emerged between the groups (p = 0.982). Some students missed one of the three test sessions due to school events, illness, or technical glitches, causing sample size variations (see Table 1).

Table 1
Sample Size for Each Test Session

	Pretest	Posttest	Delayed Posttest
Treatment Group	24	28	23
Control Group	18	23	24
Total	42	51	47

Additionally, 25 adult native English speakers from Australia (M_{age} = 28.42, *SD* = 11.11) were recruited via the online bulletin of the University of Melbourne for baseline data collection. All participants but one had, at minimum, a bachelor's degree or were pursuing one. The lone exception held only a high school diploma and was not enrolled in further education.

Reading Training

Chunk Reading Training (CRT). For the treatment group, a CRT involving 600 sentences was developed. The sentences followed either a single- or a multiple-clause structure. Sentences following the multiple-clause structure included conjunctions (e.g., *and, but*) or subordinate conjunctions (e.g., *when, because*). The following five types of verb constructions were used in the CRT (see Table 2): Type A included phrasal verbs with a direct object, and Type B included a simple transitive construction with a transitive verb and a direct object. Types C and D consisted of a caused-motion construction and a ditransitive construction, respectively (Goldberg, 2019). Type E had a be-verb and an adjective/noun phrase and did not have a transitive verb or direct object; it was included for a wider variety of sentence constructions. All participants had learned these construction types in middle school.

Table 2
Target Constructions for Chunk Reading Training

Label	Form *Example*
(A) Phrasal Verb Construction	Verb + Preposition + Object *She waited for him.*
(B) Simple Transitive Verb Construction	Verb + Object *The boy was reading a book.*
(C) Caused-Motion Construction	Subject + Verb + Object + $Obique_{path}$ *She put the ball in the box.*
(D) Ditransitive Construction	Subject + Verb + Object + Object2 *She gave him something.*
(E) Additional	Be Verb + Adjective/Noun *They were kind.* *He was a student.*

Chunking methods were important for this training (see Table 3). All training sentences were chunked using the chunking methods below. Importantly, in line with Rule [2], a verb and an object were chunked together because this chunking method was expected to help Japanese EFL learners process a verb and an object as one chunk and get used to processing the [S]VO order. Sentences (5) to (8) present examples of chunking methods.

Table 3
Chunking Methods for Chunk Reading Training

Fundamental Chunking Method
[1] "Chunk" is defined as a unit consisting of one semantic head and function/modifying words (Abney, 1991).
Other Chunking Methods
[2] A (phrasal) verb and object phrase are included in one chunk.
[3] An auxiliary verb (e.g., *can*, *must*) is attached to the verb phrase.
[4] A (subordinate) conjunction is marked as one chunk (Yamashita & Ichikawa, 2010).
[5] A be-verb and adjective/noun phrase are contained in one chunk (construction Type E).

(5) *Any place can be a school / if / you / can learn something / at that place.*
(6) *My family / will have a party / at home / this Sunday / because / I / graduated from my university.*
(7) *It was Christmas / last week, / so / the students / sent a Christmas card / to their teachers / at school.*
(8) *My parents / gave me a present / for my birthday / last night.*

In addition to the linguistic chunking rules, chunking norming data were collected from 100 L1-English speakers via *Amazon Mechanical Turk* to determine how these speakers segment English sentences into smaller multiword units. During the norming task, participants segmented 60 short sentences consisting of a subject phrase + verb phrase + prepositional phrase 1 + prepositional phrase 2 (e.g., *The students sent a Christmas card to their teachers at school*) with the target constructions (15 sentences each) into smaller units that were syntactically and semantically natural by inserting slashes. The most preferred chunking methods were consistent with the linguistic chunking rules given above (see Appendix).

To form 600 training sentences, story passages were extracted from English textbooks used at Japanese high schools and from a Japanese English proficiency test called *Eiken* (Grades 3 and pre-2; Story Sessions) (Eiken Foundation of Japan, n.d.). Some sentences were modified by the author to control the sentence structures and constructions. It was difficult to prepare an equal number of sentences for the four construction types (A to D) in Story Sessions, so original sentences were generated (non-Story Sessions). Each construction type appeared roughly 150 times (± 20 times) in the CRT. A total of 10 training sessions involving 60 sentences were prepared. The training time for each session was about 15 minutes.

The CRT was programmed using Qualtrics (see Figure 1). In the training sessions, the first chunk showed up at the center of the computer screen after a starting signal (+). The first chunk was automatically replaced by the second chunk, and this sequence of events continued to the last chunk, followed by a period. Thus, participants could not reread a chunk once the next chunk replaced it. After the final chunk, a full sentence was presented for four seconds so that the learners could check the information they missed when reading the presented chunks. The presenting time was set at 0.5 seconds per syllable, but words with one syllable (e.g., *she*, *they*) were presented for 1.5 seconds for word identification. True/false questions about the previously presented sentence were shown in Japanese after every three sentences. For each question, the participants had to click a true (O) or false (X) button within 10 seconds. When no question was shown, they clicked on “No answer” (in Japanese) to proceed to the next sentence.

Figure 1
Example of Chunk Reading Training

Reading training in block format. The participants in the control group received reading tasks extracted from *Eiken* (Grades 3 and pre-2). This training was also programmed using Qualtrics. In each training session, two passages of 250–300 words, each accompanied by four comprehensive questions, were provided in the traditional block format. In the format, sentences were displayed in paragraphs on the computer screen, instead of being segmented into smaller multiword units. To ensure that the treatment and control groups both worked on the training sessions for the same length of time, the training time for the control group was adjusted according to the CRT.

Reading Span Test

An automated RST (Farmer et al, 2017) was administered as the pretest, posttest, and delayed posttest in this study to measure the participants’ verbal WM. The fundamental method was based on Harrington and Sawyer’s (1992) method for L2/EFL learners of English. Participants were required to read unrelated sentences presented at the center of a computer screen within 10 seconds and judge their grammaticality. They were also required to remember the final word of each sentence and were prompted to retrieve these words after reading a predetermined number of sentences. All final words were singular nouns, and no words were repeated during a test session.

Sentences were generated using the four constructions previously mentioned (A, B, C, and D); they were simple, active, and 10–14 syllables in length. Since the four construction types were used to decide on the chunking methods more rigorously and not to analyze learners' processing tendencies, the construction type was not an independent variable for data analysis. The following sentences present examples of stimuli sentences.

(a) *I have waited for my friends at the station.* (Grammatical)
(b) *The driver sent [the to woman the] airport.* (Ungrammatical)

While example (a) is a grammatical sentence, example (b) is ungrammatical. In ungrammatical sentences, 3–4 preterminal words were reversed (i.e., from *the woman to the* to *the to woman the*). The study participants judged the grammaticality of each sentence and pressed the F key for grammatical sentences and the J key for ungrammatical sentences.

Sentences were presented in two-, three-, four-, and five-sentence conditions, and each condition involved three WM sessions, resulting in a total of 42 sentences (21 = grammatical). Before each sentence appeared, the "+" signal was presented on a computer screen for 500 ms. Subsequently, if the participant could not judge the grammaticality of the presented sentence within 10 seconds, the sentence would automatically disappear from the screen. After reading a certain number of sentences (i.e., 2–5 sentences), the participants were prompted to type as many words as possible in the blanks provided. However, they were instructed not to begin their recall with the last word presented (Yoshimura et al, 2021). Minor spelling errors were ignored.

Three tests were prepared—that is, a pretest, a posttest, and a delayed posttest—and they were counterbalanced to avoid any testing effects. The last words of the sentences were the same across the three tests to control for word frequency and syllables. However, the order of these last words was completely randomized across the tests, and apart from the last words, all sentences differed. There were no significant differences between the number of sentence syllables in each test ($p = 0.763$). The RST was programmed using PsyToolkit (Stoet, 2010, 2017).

Procedure

Figure 2 illustrates the procedure used for the intervention study. On the first day, the Japanese participants gathered in a computer lab at school and took the RST individually. The CRT and the reading training in block format began two weeks later as homework. As the participants already had their own accounts on an online learning platform (Moodle) for their regular classes, the link to each online training session was uploaded to Moodle every two days. The students took the posttest one day after the last training session and the delayed posttest two weeks after the posttest. The author kept track of the completion of the training programs, and no one underwent the training session more than once. Data collection from the L1 speakers was conducted online. Each participant participated individually in a face-to-face online session with the author.

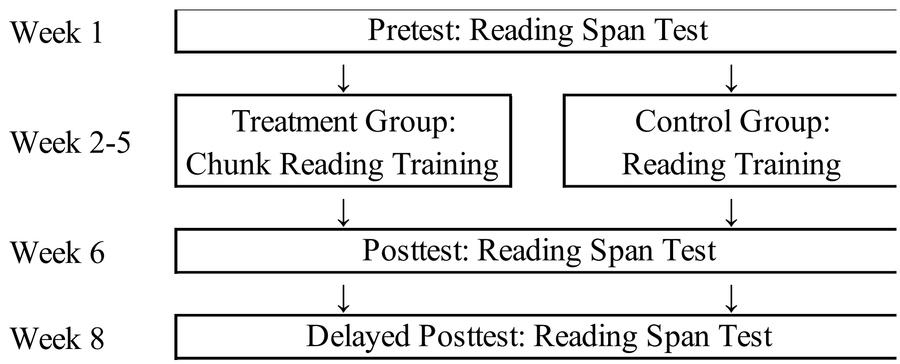

Figure 2
Procedure of Experiment

Data Analysis

The effects of the training sessions were compared over two time periods: the pretest–posttest and pretest–delayed posttest periods. For the GJT, correct responses were coded 1, and incorrect and timed-out responses were coded 0. Similarly, for the WRT, correctly and incorrectly recalled words were coded 1 and 0, respectively. Accuracy and recalled words were first analyzed separately. The accuracy and recall rates (%) were then converted into standardized *z*-scores, and the composite *z*-scores of the two components were calculated as the final reading spans (Alptekin & Erçetin, 2010). Binomial generalized linear mixed-effects (LME) modeling was conducted to answer RQ1 and to analyze word recall rates (i.e., binomial data), while an LME model was fitted to the composite *z*-score (i.e., continuous data) to answer RQ2. LME modeling is robust to missing data in a cross-sectional study (Linck & Cunnings, 2015). Prior to each analysis, the statistical assumptions were verified (e.g., the normality and homoscedasticity of residuals).

All the statistical analyses were performed in R version 4.1.0 (R Core Team, 2021) with the lme4 package version 1.1.27.1 (Bates et al., 2015). The independent variable "Group" (i.e., treatment vs. control) was sum coded, and "Session" (pretest, posttest, and delayed posttest) was treatment coded. Alpha was set at 0.05 for all analyses. For accuracy rate and recall rate (i.e., logistic regression models), odds ratios (OR) were calculated as effect sizes. Odds ratios greater than 3 or less than 0.33 were considered to represent strong effects (Haddock et al.,1998). For the composite *z*-score, *d* values were calculated using the *eff_size* function from the emmeans package version 1.7.5 (Lenth, 2022). For within-group contrasts, *d* values of 0.60, 1.00, and 1.40 were considered small, medium, and large, respectively (Plonsky & Oswald, 2014). All post-hoc tests were conducted using the estimated marginal means function from the emmeans package (Lenth, 2022) with a Tukey-adjusted *p*-value.

Results

Baseline Data

Table 5 presents a descriptive summary of the L1 speakers, which were used as baseline data. Therefore, a composite *z*-score was not calculated, and statistical analyses were not performed. The baseline data were relatively higher than those in previous studies with L1-English speakers (e.g., Farmer et al, 2017) because the task difficulty was adjusted for the Japanese participants.

Table 5
Descriptive Summary of Reading Span Test for L1 English Speakers (M [SD])

	Accuracy Rate (%)	Recall Rate (%)
L1 Speakers	0.94 (0.24)	0.88 (0.32)

Grammaticality Judgment Task

RQ 1. RQ 1 focused on whether CRT develops Japanese EFL learners' implicit and online syntactic processing skills, data from the GJT were analyzed separately from the WRT. Table 6 provides a descriptive summary of the accuracy rates, and Figure 3 visualizes the results. In the pretest, the accuracy rate of the treatment group was lower than that of the control group, but the recall rate of the former was higher (see the following section); thus, it is likely that both groups experienced some trade-offs between syntactic processing and the storage of sentence-final words.

Table 6
Descriptive Summary of Grammaticality Judgment Task (M [SD])

	Accuracy Rate (%)		
	Pretest	Posttest	Delayed Posttest
Treatment Group	0.54 (*0.50*)	0.64 (*0.48*)	0.64 (*0.48*)
Control Group	0.62 (*0.49*)	0.61 (*0.49*)	0.65 (*0.48*)

A generalized LME model was fitted with Session and Group as the categorical variables and with an interaction term. The pretest session was set as the reference level for Session. The model also included random by-subject and by-item intercepts but not slopes due to the small sample size; this was in line with Saito et al's (2022) research. Proficiency (i.e., the learners' midterm exam score) was included as a covariate.

Session was found to have a significant main effect. The overall accuracy rate, averaged across the treatment and control groups, increased from the pretest ($M = 0.57$, $SD = 0.49$) to the posttest ($M = 0.62$, $SD = 0.48$) ($\beta = 0.27$, $SE = 0.07$, $z = 3.78$, $p = 0.0002$, OR = 1.32) and from the pretest to the delayed posttest ($M = 0.64$, $SD = 0.48$) ($\beta = 0.37$, $SE = 0.08$, $z = 4.89$, $p < 0.0001$, OR = 1.45). The fixed effect of Group was significant ($\beta = 0.17$, $SE = 0.08$, $z = 2.09$, $p = 0.036$). However, since Session was treatment coded prior to analysis, Group represents a simple effect, not a main effect corresponding to an analysis of variance (ANOVA) model (Brehm & Alday, 2022).

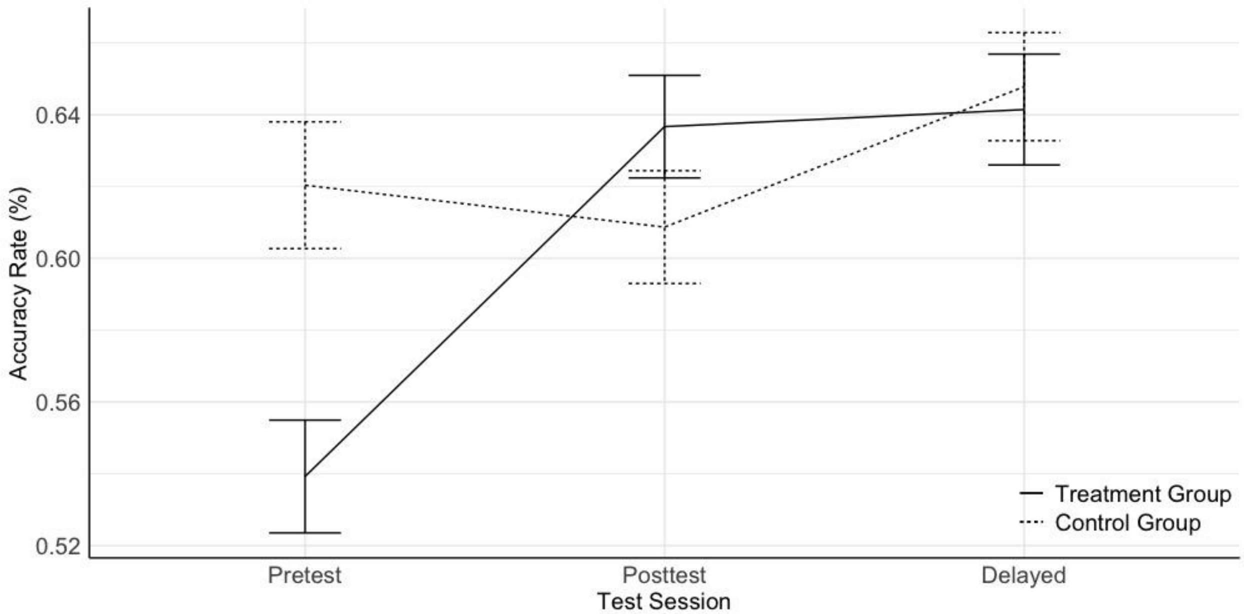

Figure 3
Accuracy Rate for Grammaticality Judgement Task (%)

A post-hoc comparison did not indicate a significant main effect for Group (β = 0.11, *SE* = 0.14, *z* = 0.76, *p* = 0.445). As a covariate, Proficiency reached significance (β = 0.01, *SE* = 0.00, *z* = 2.76, *p* = 0.006), suggesting that individuals with higher midterm exam scores achieved higher accuracy rates. A significant interaction between Session and Group was found in the posttest (β = -0.22, *SE* = 0.07, *z* = -3.08, *p* = 0.002), and a marginally significant interaction was seen in the delayed posttest (β = -0.14, *SE* = 0.07, *z* = -1.84, *p* = 0.066). To further explore the significant interaction, a post-hoc test was performed, and it revealed that accuracy rate significantly improved from the pretest to the posttest for the treatment group (β = 0.50, *SE* = 0.10, *z* = 5.11, *p* < 0.0001, OR = 1.65) but not for the control group (β = -0.05, *SE* = 0.11, *z* = -0.48, *p* = 0.997, OR = 0.95). However, there were no significant differences between the two groups in any of the test sessions (*ps* > 0.3).

Word Recall Task

The data from the WRT were analyzed; Table 7 provides the descriptive summary of the word recall rate, and Figure 4 visually depicts the results. As mentioned above, the control group achieved a lower recall rate than the treatment group did in the pretest. Both groups may have experienced some trade-offs between syntactic processing and the storage of sentence-final words.

Table 7
Descriptive Summary of Word Recall Rate (M [SD])

	Recall Rate (%)		
	Pretest	Posttest	Delayed Posttest
Treatment Group	0.62 (0.49)	0.70 (0.46)	0.71 (0.45)
Control Group	0.54 (0.50)	0.63 (0.48)	0.64 (0.48)

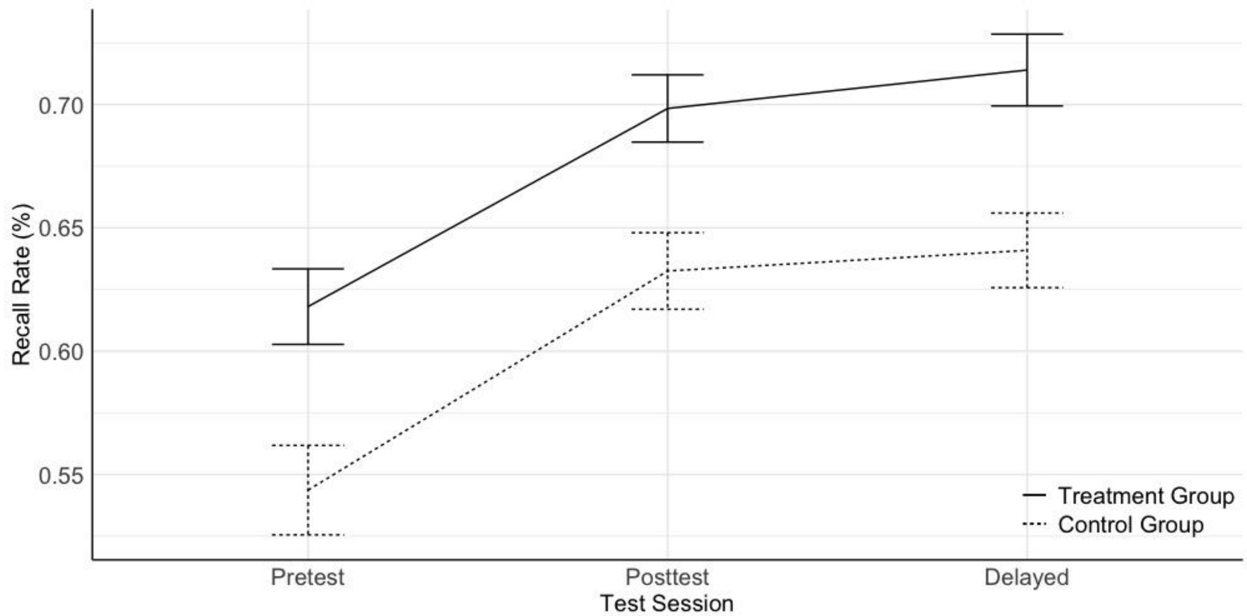

Figure 4
Word Recall Rate (%)

Generalized LME modeling was conducted. The formula was the same as that used for the GJT. The model yielded a main effect for Session, indicating that the overall recall rate, averaged across the two groups, increased from the pretest (M = 0.59, SD = 0.49) to the posttest (M = 0.67, SD = 0.47) (β = 0.45, SE = 0.08, z = 5.68, p < 0.0001, OR = 1.56) and from the pretest to the delayed posttest (M = 0.68, SD = 0.47) (β = -0.58, SE = 0.08, z = 7.03, p < 0.0001, OR = 1.78). A post-hoc comparison did not reveal a significant main effect for Group (β = -0.26, SE = 0.22, z = -1.20, p = 0.228). Further, neither the covariate (β = 0.00, SE = 0.01, z = 0.35, p = 0.727) nor the interactions (pretest–posttest: β = -0.13, SE = 0.08, z = -0.55, p = 0.581, pretest–delayed posttest: β = -0.13, SE = 0.08, z = -1.57, p = 0.117) reached significance.

Reading Spans (Composite Z-Scores)

RQ 2. RQ 2 addressed whether CRT expands Japanese EFL learners' verbal WM. Composite z-scores were calculated based on the accuracy and word recall rates, representing the final reading spans. The composite z-scores were collapsed and calculated over the three test sessions so that they could be analyzed on the same scale. Table 8 provides a descriptive summary of the composite z-scores, and Figure 5 visually depicts the results.

Table 8
Descriptive Summary of Composite Z-Scores (M [SD])

	Pretest	Posttest	Delayed Posttest
Treatment Group	-0.35 (0.50)	0.18 (0.78)	0.30 (0.61)
Control Group	-0.27 (0.74)	-0.06 (0.68)	0.11 (0.67)

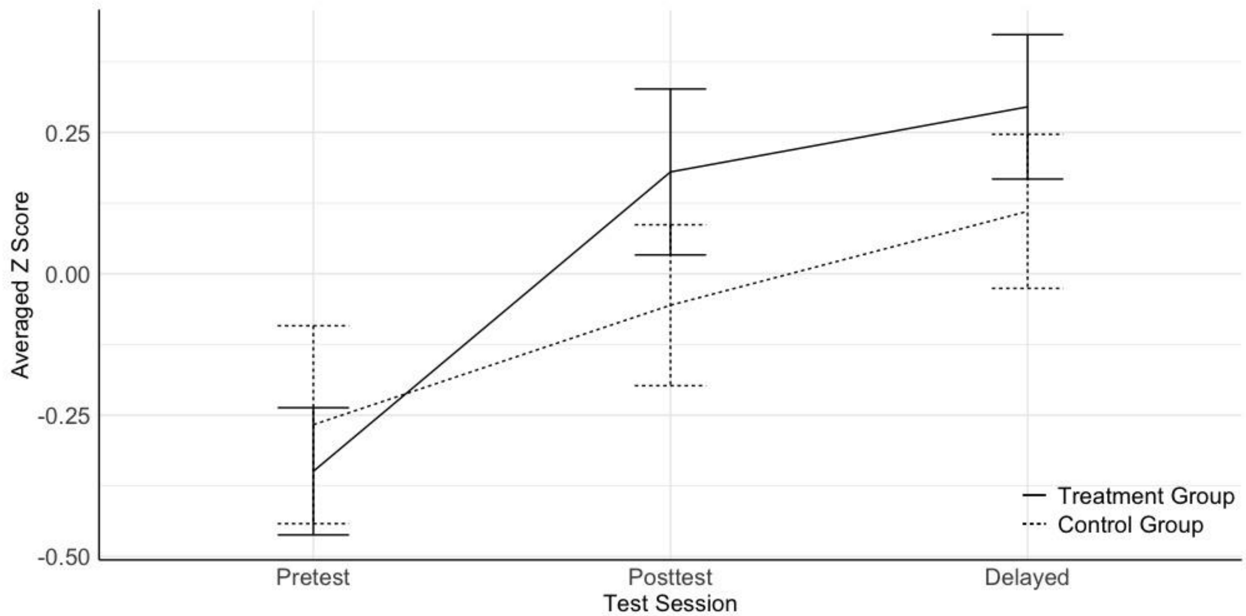

Figure 5
Composite Z-Scores of Reading Span Test

An LME model was fitted with Session and Group as categorical variables and Proficiency as a covariate. Since all accuracy and word recall rates were averaged and converted into *z*-scores, the model included only the by-subject random intercept. The overall composite *z*-score, averaged across the two groups, improved significantly from the pretest (M = -0.31, SD = 0.63) to the posttest (M = 0.07, SD = 0.74) (β = 0.44, SE = 0.09, t = 4.99, p < 0.0001, d = 1.09) and from the pretest to the delayed posttest (M = 0.20, SD = 0.64) (β = 0.59, SE = 0.09, t = 5.56, p < 0.0001, d = 1.47).

A post-hoc comparison did not reveal a significant main effect for Group (β = -0.05, SE = 0.16, t = -0.33, p = 0.742). An interaction between Session and Group was significant in the posttest (β = -0.18, SE = 0.09, t = -2.09, p = 0.040) and marginally significant in the delayed posttest (β = -0.16, SE = 0.09, t = -1.75, p = 0.084).

A post-hoc test revealed that the *z*-scores significantly improved from the pretest to the posttest for the treatment group (β = 0.62, SE = 0.12, z = 5.20, p < 0.0001, d = 1.54) but not for the control group (β = 0.26, SE = 0.14, z = 1.89, p = 0.412, d = 0.63). There were no significant differences between the two groups in all the test sessions (*ps* > 0.9).

Discussion

The purpose of the present study was to investigate whether CRT facilitates the growth of Japanese EFL learners' implicit online syntactic processing skills and verbal WM. The Chunk-and-Pass model posits that a language user must chunk language inputs into large units and recode them into abstract representations (Christiansen & Chater, 2016). From the Chunk-and-Pass perspective, language acquisition can be conceptualized as the process of acquiring the ability to appropriately chunk language inputs (Christiansen & Chater, 2016). This is based on the development of schematic representations of linguistic constructions (Goldberg, 2019), which enable learners to rapidly chunk incoming language (Christiansen & Chater, 2016). Drawing from the theoretical framework posited by the model, the CRT in the present study was aimed at implicitly teaching the participants how to process English sentences by presenting multiword chunks that are syntactically and semantically complete.

With regard to the RST, both groups exhibited a certain degree of trade-offs between the GJT and the WRT in the pretest stage: the treatment group obtained lower accuracy rates on the GJT and higher word recall rates on the WRT than the control group, while the control group showed the opposite trend. This indicates that the Japanese participants allocated more cognitive resources to syntactic processing or word storage. However, in the posttest, the treatment group had a significantly improved accuracy rate on the GJT, while the control group did not.

This finding is consistent with a previous L2 learning study by Grey et al (2015), who used a GJT similar to the one in the present study and reported that L1-English learners of advanced L2 Spanish at the college level had improved accuracy rates on the GJT after five weeks of studying abroad (12.89% increase). Specifically, the participants were immersed in a Spanish-speaking environment for a period of five weeks, with academic activities consuming their time from 10 a.m. to 9 p.m. for four days a week. Unlike Grey et al (2015), the present study showed that the Japanese EFL learners could improve their GJT performance (i.e., 10% increase) with only 15 minutes of CRT for three days a week across four weeks. Although there were no statistically significant differences between the groups across the test sessions, the results indicate that CRT has the potential to positively affect the development of learners' syntactic processing, without requiring an extended workload.

A chunk reading strategy can reduce learners' cognitive loads during text processing (Tate et al, 2019), as the meaningful unit-based segmentation of information facilitates the recognition of information boundaries (in this case, syntactic and semantic boundaries) (Wouters et al, 2008). In addition, segmented texts can draw learners' attention toward syntactic structures that are usually ignored when they read sentences in block format (Park, 2018). It can thus be concluded that the treatment group's cognitive load during text processing was reduced due to the participants' exposure to syntactically and semantically complete multiword units, allowing them to focus on processing phrasal constructions in the English word order. Most of the reversed words in the sentence stimuli included in the RST were part of a verb and a prepositional phrase, as follows.

(9) Stimulus sentence: **The woman was singing the by song a river.*

(10) Original sentence: *The woman / was singing a song / by the river.*

In (9), part of the verb phrase (*a song*) and part of the prepositional phrase (*by the*) have been reversed. Consequently, it was inferred that the treatment group exhibited a greater level of accuracy in identifying the ungrammaticality of an uncanonical word order after the main verb in the posttest, reflecting the development of their schematic knowledge on the English (S)VO word order and the word order of prepositional phrases (i.e., preposition + noun phrase [*in the river*]) through CRT. In short, CRT was proven to have the potential to cultivate Japanese EFL learners' abstract syntactic knowledge and thereby enable them to process syntactic information accurately during sentence processing. This finding provides partial support for the hypothesis underlying the Chunk-and-Pass model, which is that language acquisition is about learning how to process a language and is driven by language processing experiences (Christiansen & Chater, 2016).

The WRT results revealed an overall improvement in the recall rates across the test sessions for both the treatment and control groups. Nonetheless, the lack of a significant interaction between Session and Group suggests that CRT had no impact on the storage capacity of the treatment group. Psycholinguistic studies have shown that as the efficiency of individual processing of a target language is more automatized, the number of items stored during an RST increases (Farmer et al, 2017; Van den Noort et al, 2006). Therefore, it is possible that the CRT implemented in the present study did not improve the treatment group's syntactic processing efficiency enough to expand their memory capacity levels during the RST.

The composite *z*-scores of the GJT and the WRT were calculated in accordance with psycholinguistic research (Alptekin & Erçetin, 2010) to obtain precise estimates of the participants' reading spans. The results showed that only the treatment group experienced a significant increase in the posttest composite *z*-score compared to the pretest score. Although the treatment group did not outperform the control group significantly after treatment, this finding highlights a trend toward improved verbal WM in the treatment group following the CRT intervention. It is evident that the improved syntactic processing skills facilitated by CRT contributed to the increased reading spans of the participants in the treatment group. The interaction between Session and Group was marginally significant for the delayed posttest, although it did not reach significance. This may be attributed to the fact that the control group's *z*-scores also improved in the delayed posttest.

As discussed above, this study did not reveal any statistical differences between the groups in the measured components, suggesting that the control group also progressed throughout the testing phases, especially in the delayed posttest. One possible interpretation for these results is that as observed in previous cross-sectional studies involving an RST (e.g., Wells et al, 2009), the control group demonstrated increasing familiarity with the test material over successive test sessions, resulting in better scores than in the pretest. In fact, the control group did not show a significant improvement in the GJT and *z*-scores from the pretest to the posttest.

This result suggests that the impact of the reading practice may have been negligible in comparison to CRT. Another possibility is that the control group's improvements were influenced by additional exposure to English outside the classroom. Recall that the length of training for this group was matched to that of the treatment group. Given that the length of exposure to the target language is a critical factor in L2 success (Ortega, 2009), it is not surprising that the control group also experienced some benefit from the additional reading practice (Malakowsky, 2023). However, even after considering the aforementioned possibilities, the marginally significant interactions between Session and Group in the delayed posttest of the GJT and the composite *z*-scores indicate that the effects of CRT were present to some extent in the RST for an extended period.

Educational Implications

Although the impact of CRT was somewhat restricted in this study due to the short duration of the program, the findings suggest that CRT can enhance EFL learners' online syntactic processing skills and verbal WM, potentially bolstering their reading comprehension in later acquisition stages (Peng et al, 2018). Recent reading research has highlighted the value of additional fluency-centric training (e.g., extensive reading) in EFL settings such as Japan, where the grammar-translation teaching method has long been favored (Sakurai, 2015). However, many of these training methods overlook the importance of nurturing fundamental language processing skills, including syntactic processing. In fact, it has been suggested that researchers seek teaching strategies for the development of EFL learners' syntactic processing abilities (e.g., Park, 2018). According to usage-based theories (e.g., Christiansen & Chater, 2016; Goldberg, 2019; Tomasello, 2003), it should be possible for EFL learners to strengthen their syntactic processing skills with ample language exposure. However, time constraints in EFL environments underscore the need for complementary teaching tools that address these skills. In this context, the present research highlighted the potential role of CRT in refining EFL learners' syntactic processing skills.

In pragmatic terms, English educators can weave CRT into their curricula and synergize them with diverse teaching methods, regardless of the teaching environment. Recent intervention studies grounded in construction grammar have shown that the combination of exposure to linguistic structures and access to explicit instructions is more effective at fostering linguistic construction acquisition than exposure alone (Holme, 2010; Rah & Kim, 2018). Thus, CRT may be able to supplement syntactic processing training, especially when paired with explicit instructions, whether integrated into the classroom or assigned as extracurricular work; the participants in this study completed the treatment through homework.

Limitations and Conclusion

This study hints at CRT's potential to enhance the syntactic processing skills and verbal WM of Japanese EFL learners; however, certain limitations exist. The training duration was shorter in this study than in prior CRT research (e.g., Tate et al, 2019; Walker et al, 2005). L2 acquisition literature emphasizes that the duration of exposure to the target language is a pivotal factor affecting L2 success (Ortega, 2009). Thus, the brief duration of CRT in the present study may have limited its instructional impact. For instance, no significant between-group differences emerged among the measured components, and although word recall rates increased overall, the interaction between Session and Group was not statistically significant. The GJT results indicate that extended CRT can boost the syntactic processing of Japanese EFL learners, thus freeing up WM cognitive resources and potentially raising recall rates.

Last but not least, as previously cautioned, the CRT program in this study was manipulated by the author to control for construction type and to make the sentences more chunkable. It should be noted that the present study evidenced that CRT has the potential to affect the development of learners' syntactic processing skills and verbal WMC; however, the aim of the study was not to compare CRT with other text formats used in existing materials for Japanese English classes. Therefore, CRT can be employed as a *supplementary pedagogical tool* in conjunction with other instructional teaching strategies, as discussed above.

Despite the considerable reactions to the Chunk-and-Pass model across various disciplines in the language sciences, empirical and pedagogical investigations involving the model remain scarce. Further research is necessary to investigate the suitability of the model's use in L2 learning contexts, considering the above limitations.

Acknowledgements

This research was taken from part of the author's Ph.D. dissertation. The author greatly appreciates the guidance and support from the dissertation committee members, which include Helen Zhao, Carsten Roever, Barbara Kelly, and Jonathon Lum. The author also thanks the participants and teachers at the Japanese secondary school for participating and supporting this research. Special thanks go to Iroha Sakata for her valuable suggestions on Japanese translations used in the test materials. This research was supported by Melbourne Research Scholarship and *Language Learning* Doctoral Dissertation Grant Award.

Ethical Approval

The present research received ethics approval from the University of Melbourne (ethics ID: 2021-22477-23964-5).

Reference

Abney, S. P. (1991). Parsing by chunks. In R. C. Berwick, S. P. Abney, & C. Tenny (Eds.), *Principle-based parsing* (pp. 257-278). Springer.

Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, *33*(2), 206-219. https://doi.org/10.1111/j.1467-9817.2009.01412.x

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *GA Bower (Ed.). Psychology of learning and motivation: Advances in research and theory* (pp. 47-89). Academic press.

Bates, C. T., & Maechler, M., Bolker, B., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, *125*, 104334. https://doi.org/10.1016/j.jml.2022.104334
Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating evolution, processing, and processing*. MIT Press.
Clark, K. B. (2018). Natural chunk-and-pass language processing: Just another joint source-channel coding model?. *Communicative & Integrative Biology*, *11*(2), 1-2. https://doi.org/10.1080/19420889.2018.1445899
Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Science, 24*(1), 87-114. https://doi.org/10.1017/S0140525X01003922
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466. https://doi.org/10.1016/S0022-5371(80)90312-6
Diessel, H. (2015). Usage-based construction grammar. In E. Dabrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 295-321). De Gruyter. https://doi.org/10.1515/9783110292022-015
Eiken Foundation of Japan (n.d.). *Downloads*. Retrieved March 1, 2022, from https://www.eiken.or.jp/eiken/en/downloads/
Farmer, T. A., Fine, A. B., Misyak, J. B., & Christiansen, M. H. (2017). Reading span task performance, linguistic experience, and the processing of unexpected syntactic events. *Quarterly Journal of Experimental Psychology*, *70*(3), 413-433. https://doi.org/10.1080%2F17470218.2015.1131310
Fender, M. (2003). English word recognition and word integration skills of native Arabic-and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics, 24*(2), 289-315. https://doi.org/10.1017/S014271640300016X
Gao, Y., & Wu, Y. (2023). The influences of working memory updating on word association effects and thematic role assignment during sentence processing. *Neuropsychologia*, *184*, 108547. https://doi.org/10.1016/j.neuropsychologia.2023.108547
Goldberg, A. E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton University Press.
Grey, S., Cox, J. G., Serafini, E. J., & Sanz, C. (2015). The role of individual differences in the study abroad context: Cognitive capacity and language development during short-term intensive language exposure. *The Modern Language Journal*, *99*(1), 137-157. https://doi.org/10.1111/modl.12190
Guo, R., & Ellis, N. C. (2021). Language usage and second language morphosyntax: Effects of availability, reliability, and formulaicity. *Frontiers in psychology*, *12*, 582259. https://doi.org/10.3389/fpsyg.2021.582259
Haddock, C. K., Rindskopf, D., & Shadish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: a primer on methods and issues. *Psychological Methods*, *3*(3), 339-353. https://psycnet.apa.org/doi/10.1037/1082-989X.3.3.339
Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*(1), 25-38. https://doi.org/10.1017/S0272263100010457
Holme, R. (2010). A construction grammar for the classroom. *IRAL, 48*(4), 355-377. https://doi.org/10.1515/iral.2010.015

Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2020). Meaningfulness beats frequency in multiword chunk processing. *Cognitive Science*, *44*(10), e12885. https://doi.org/10.1111/cogs.12885

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*(1), 122-149. https://psycnet.apa.org/doi/10.1037/0033-295X.99.1.122

Lenth, V. R. (2022). emmeans: Estimated marginal means, aka least-squares means. (Version 1.7.5.) [Computer software]. Retrieved from https://CRAN.R-project.org/package=emmeans

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, *65*(S1), 185-207. https://doi.org/10.1111/lang.12117

Lupyan, G., & Christiansen, M. H. (2002). Case, word order, and language learnability: Insights from connectionist modeling. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 596-601). Lawrence Erlbaum Associates. https://escholarship.org/content/qt8nf95595/qt8nf95595.pdf

MacDonald, C. M. & Christiansen, M. H. (2002). Reassessing working memory: a reply to Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review, 109*(1), 35-54. https://psycnet.apa.org/doi/10.1037/0033-295X.109.1.35

McMaster, K. L., van den Broek, P., Espin, C. A., Pinto, V., Janda, B., Lam, E., ... & van Boekel, M. (2015). Developing a reading comprehension intervention: Translating cognitive theory to educational practice. *Contemporary Educational Psychology*, *40*, 28-40. https://doi.org/10.1016/j.cedpsych.2014.04.001

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97. https://psycnet.apa.org/doi/10.1037/h0043158

Malakowsky, D. (2023). A Modified Extensive Reading and Repeated Reading Intervention with Adult ESL Students. *Reading in a Foreign Language*, *35*(1), 72–99. https://hdl.handle.net/10125/67439

Ortega, L. (2009). *Understanding second language acquisition.* Hodder.

Park, Y. (2018). Syntactic enhancement: Bootstrapping for second language reading. *Journal of Cognitive Science*, *18*(4), 473-509.

Park, Y. & Warschauer, M. (2016). Syntactic enhancement and second language literacy: An experimental study. *Language Learning & Technology, 20*(3), 180-199. http://dx.doi.org/10125/44488

Park, Y., Xu, Y., Collins, P., Farkas, G., & Warschauer, M. (2019). Scaffolding learning of language structures with visual-syntactic text formatting. *British Journal of Educational Technology*, *50*(4), 1896-1912. https://doi.org/10.1111/bjet.12689

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological bulletin*, *144*(1), 48-76. https://psycnet.apa.org/doi/10.1037/bul0000124

Perez, M. M. (2020). Incidental vocabulary learning through viewing video: The role of vocabulary knowledge and working memory. *Studies in Second Language Acquisition*, *42*(4), 749-773. https://doi.org/10.1017/S0272263119000706

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language learning*, *64*(4), 878-912. https://doi.org/10.1111/lang.12079
R Core Team. (2021). R: A language and environment for statistical computing (Version 4.1.0) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing https://www.R-project.org/
Rah, Y., & Kim, H. (2018). Construction-based approach to teaching the English resultative construction to Korean EFL learners. *System, 72*(1), 1-12. https://doi.org/10.1016/j.system.2017.09.027
Saito, K., Hanzawa, K., Petrova, K., Kachlicka, M., Suzukida, Y., & Tierney, A. (2022). Incidental and multimodal high variability phonetic training: Potential, limits, and future directions. *Language Learning*, *72*(4), 1049-1091. https://doi.org/10.1111/lang.12503
Sakurai, N. (2015). The influence of translation on reading amount, proficiency, and speed in extensive reading. *Reading in a Foreign Language*, *27*(1), 96-112. http://hdl.handle.net/10125/66702
Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior research methods*, *42*(4), 1096-1104. https://doi.org/10.3758/BRM.42.4.1096
Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24-31. https://doi.org/10.1177/0098628316677643
Tate, T. P., Collins, P., Xu, Y., Yau, J. C., Krishnan, J., Prado, Y., ... & Warschauer, M. (2019). Visual-syntactic text format: Improving adolescent literacy. *Scientific Studies of Reading*, *23*(4), 287-304. https://doi.org/10.1080/10888438.2018.1561700
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge. Harvard University Press.
Van den Noort, M. W., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*(4), 289-296. https://psycnet.apa.org/doi/10.1027/1016-9040.11.4.289
Walker, S., Schloss, P., Fletcher, C. R., Vogel, C. A., & Walker, R. C. (2005). Visual-syntactic text formatting: A new method to enhance online reading. *Reading Online*, *8*(6), 1096-1232.
Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250-271. https://doi.org/10.1016/j.cogpsych.2008.08.002
Wouters, P., Paas, F., & van Merriënboer, J. J. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, *78*(3), 645-675. https://doi.org/10.3102/0034654308320320
Yamashita, J., & Ichikawa, S. (2010). Examining reading fluency in a foreign language: Effects of text segmentation on L2 readers. *Reading in a Foreign Language*, *22*(2), 263-283. http://hdl.handle.net/10125/66841
Yoshimura, T., Osaka, M., Osawa, A., & Maeshima, S. (2021). The classical backward digit span task detects changes in working memory but is unsuitable for classifying the severity of dementia. *Applied Neuropsychology: Adult*, 1-7. https://doi.org/10.1080/23279095.2021.1961774

Zhang, X., & Dong, X. (2019). Input frequency and construction interference interactions in L2 development. *Second Language Research*, *35*(4), 505-527. https://doi.org/10.1177/0267658318791651

Appendix

Summary of Norming Data for Chunking Methods

		SP	VP	PP 1	PP2
A	Structure	NP	V + Prep + Obj.	Prep. + NP	Prep. + NP
	Example	*The teacher*	*waited for the student*	*at the station*	*for one hour*
	Ratio	99%	62%	87%	100%

		SP	VP	PP 1	PP2
B	Structure	NP	V + Obj.	Prep. + NP	Prep. + NP
	Example	*The boy*	*ate lunch*	*at the restaurant*	*with his family*
	Ratio	99%	78%	99%	89%

		SP	VP	$\text{Oblique}_{\text{path}}$	PP2
C	Structure	NP	V + Obj.	Prep. + NP	Prep. + NP
	Example	*The woman*	*sent a letter*	*to her boyfriend*	*last week*
	Ratio	99%	75%	82%	93%

		SP	VP	PP 1	PP2
D	Structure	NP	V + Obj. 1 + Obj. 2	Prep. + NP	Prep. + NP
	Example	*The man*	*gave her a present*	*for her birthday*	*last inght*
	Ratio	99%	49%	99%	84%

Note. SP: subject phrase, VP: verb phrase, PP: prepositional phrase, NP: noun phrase, V: verb, Obj.: object, Prep: preposition. The selected structures were the most preferred chunking methods by 100 L1-English speakers. For construction D (ditransitive construction), the second preferred structure was "/V + Obj. 1 / Obj. 2/ (e.g., /*gave her / a present/*)" with about 21%. Since the present CRT aimed to teach the L1-Japanese learners of English to process a verb and objects in one chunk, the structure in the table (V + Obj. 1 + Obj. 2) was applied although the ratio was relatively low (49%).

About the Author

Takumi Kosaka is a Ph.D. candidate at the University of Melbourne. His research interests include psycholinguistics and TESOL. He has experience teaching English at Japanese secondary schools and at a community college in the United States.
E-mail: tkosaka@student.unimelb.edu.au