日本語の文理解における，年齢，語順および文型による影響 ――韓国語を扱ったSung et al.（2017）とSung（2015）の追試――

Effects of Age, Word Order, and Sentence Types on Japanese Sentence Comprehension: A Replication Study of Sung et al. (2017) and Sung (2015) on Korean

有井　巴 ARII, Tomoe

● 横浜市立大学
Yokohama City University

李　勝勲 LEE, Seunghun J.

● 国際基督教大学，ベンダ大学
International Christian University/ University of Venda

吉田　智行 YOSHIDA, Tomoyuki

● 国際基督教大学
International Christian University

成　知垠 SUNG, Jee Eun

● 梨花女子大学
Ewha Womans University

Keywords 文理解，高齢化，日本語，語順，文型
sentence comprehension, aging, Japanese, word order, sentence types

ABSTRACT

本追試研究は，Sung et al.（2017）とSung（2015）が韓国語の文理解研究に用いた方法を用い，日本語の文理解を調べた。Sung et al.（2017）は，年齢，語順，およびワーキングメモリーが文理解に影響することを発見した。Sung（2015）は，高齢者の文理解に焦点をあて，高齢者が文型に影響を受けることを発見

した。本研究は日本語を母語とする若者と高齢者を対象とし，年齢，語順，および文型が文理解に与える影響を調べた。文理解タスクには，次の三種の基本語順文と非基本語順文を用いた：(a) 二項動詞を含む能動文，(b) 三項動詞を含む能動文，(c) 受身文。これより，韓国語と同様，日本語の文理解にも年齢と語順が有意に影響することが分かった。また，年齢と語順に交互作用が見られた。これは，高齢者が若者より非基本語順の理解で苦労することを意味する。本研究は先行研究と異なり，日本語の文理解に態は影響しないことを発見をした。このように，本研究は追試研究に止まらず，年齢が日本語の文理解能力にもたらす変化についての新たな知見をもたらした。

This replication study examines the comprehension of Japanese sentence types using methods developed for investigating Korean sentence comprehension in Sung et al. (2017) and Sung (2015). Sung et al. (2017) found the effects of age, word order, and working memory when participants perform a sentence-comprehension task. Sung (2015) focused on the sentence comprehension of elderly adults, and found that elderly adults performed differently depending on sentence types. In our study, young and elderly adults of Japanese were recruited to examine the effects of age, word order, and sentence types on sentence comprehension. Our sentence comprehension task presented canonical and non-canonical sentences of the following three types: (a) active sentences with a two-place verb, (b) active sentences with a three-place verb, and (c) passive sentences. The Japanese participants displayed significant effects of age and word order as the Korean participants did. We also found a significant interaction between age and word order. Our results have shown that elderly adults have more difficulties with non-canonical sentences than young adults. Some results in our study deviate from the original studies: unlike the Korean studies, the voice of sentences did not show difference between active and passive voice. In sum, our study not only replicated parts of earlier studies but also contributed to better understanding the nature of the age-related changes in sentence-comprehension abilities in Japanese.

1. Introduction

Aging has been attested to relate to the decline in sentence-comprehension abilities in many previous studies (Caplan et al., 2011; DeDe et al., 2004; Kemper & Kemtes, 1999; Zurif et al., 1995). In sentence processing, working memory (WM) plays an important role by maintaining information and computing syntactic structures to interpret the correct meaning of a sentence (Just & Carpenter, 1992). Individual differences in WM have been reported to relate to sentence-comprehension abilities (e.g., Waters & Caplan, 2003). WM capacity is also well known to decline as people get older (Carpenter et al., 1995; Morris et al., 1988; G. Waters & Caplan, 2001, 2005). Waters & Caplan (2001, 2005) demonstrated that elderly adults performed significantly worse on various WM tasks than young adults. The decline of WM capacity with age is thus likely to relate to decreased sentence-comprehension abilities.

These findings came mainly from experimental studies on English participants, and replicated studies in other languages were required to better understand the general nature of the age-related changes in sentence-comprehension abilities. Sung et al. (2017) investigated Korean participants' sentence comprehension, and reported that age and WM affected sentence-comprehension abilities in this language. Korean is a verb-final language and has different basic word order (SOV) from English (SVO). Korean also exhibits considerable flexibility

in word order due to scrambling. Given that the effects of age and WM were also observed in Korean, which is typologically different from English, Sung et al. (2017) contributed much to the understanding of the general nature of the age-related decline in sentence-comprehension abilities. The purpose of the current study is to replicate the investigation of relationship between sentence comprehension and aging in Sung et al. (2017), by examining sentence comprehension in Japanese, another SOV language with flexible word order due to scrambling. While Sung et al. (2017) examined the effects of age and WM on sentence-processing and argues the role of WM in sentence-comprehension abilities, this paper focuses on the aging effect only due to our inaccessibility to test instruments measuring WM.

Sung et al. (2017) examined participants' comprehension of passive sentences in two different word orders, as shown in (1), by employing a sentence-picture-matching paradigm.

(1) a. Passive-canonical
Nolangi-ka phalangi-eykey cha-i-ta.
the Yellow-NOM the Blue-by kick-PASS-IND
'The Yellow was kicked by the Blue.'
b. Passive-noncanonical
Phalangi-eykey nolangi-ka cha-i-ta.
the Blue-by the Yellow-Nom kick-PASS-IND
'The Yellow was kicked by the Blue. (lit. By the Blue, the Yellow was kicked)'

In the study, a total number of 134 individuals (53 male and 81 female) participated. The range of the age was 21 to 86 (mean = 56, *SD* = 17). The participants were asked to point to the picture that corresponds to the target sentence. This study employed four measures of WM capacity: digits-forward (DF), digits-backward (DB), words-forward, and words-backward (DF and DB tasks from the Korean version of the Wechsler Adult Intelligence Scale (Yeom et al., 1992)). Sung et al. (2017) used a linear mixed effects model for analyzing the experimental data. They found that age negatively correlated with WM capacity and performance in the passive-sentence-comprehension task. These findings align with previous studies (Caplan et al., 2011; Kemper & Kemtes, 1999; Stine-Morrow et al., 2000; G. Waters & Caplan, 2001, 2005, 2005). Sung et al. (2017) also reported a significant interaction between age and WM; the WM effects were greater than the age effects on the task. Furthermore, the significant effects of word order were found; canonical word order (e.g., (1a)) is easier to comprehend than non-canonical word order (e.g., (1b)). Based on these findings, Sung et al. (2017) concluded that the WM effects on passive-sentence comprehension increase dramatically as people get old.

The current study examined Japanese participants to find out whether age, canonicity of word order, and sentence types (such as voice and argument numbers) have effects on the sentence comprehension in Japanese. In the same way as Korean, Japanese is an SOV language with flexible word order due to scrambling (see Saito (1985); see Nakayama (1995) and Nemoto (2017) for summary). In psycholinguistic studies, sentences with scrambled word order have been reported to be more difficult to process for young adults than their canonical counterparts (Aoshima et al., 2004; Chujo, 1983; Mazuka et al., 2002; Miyamoto & Takahashi, 2004; Tamaoka et al., 2005). Young adults can assign a correct interpretation to non-canonical sentences, but it takes longer to process them. Tamaoka et al. (2005) reported that graduate and undergraduate students showed slower reaction time when processing non-canonical sentences than canonical sentences. Such

difficulties have been attributed to *filler-gap dependency* (Aoshima et al., 2004; Chujo, 1983; Miyamoto & Takahashi, 2004). For example, in (2) below, parsers must determine whether the accusative NP in the sentence-initial position is appropriate for the object, which is typically placed immediately preceding the verb in the canonical word order. They then need to search for the original object position (i.e., the gap) to form a filler-gap dependency. Because of this extra parsing, it takes longer to process non-canonical sentences than the canonical counterparts.(i)

(2) John-o Mary-ga osita.
John-ACC Mary-NOM pushed
'Mary pushed John.'

Such difficulties in parsing non-canonical sentences have been reported only regarding Japanese young adults. To the best of our knowledge, no studies have examined Japanese elderly adults' comprehension of non-canonical sentences. Our study is thus the first attempt in Japanese to examine whether word order has effects on elderly adults' sentence comprehension and whether aging affects sentence comprehension by targeting both young and elderly adults.

While Sung et al. (2017) examined passive-sentence comprehension only, our study employed three different sentence types such as active sentences with two-place verbs, active sentences with three-place verbs, and passive sentences, following Sung (2015). The sentence examples are given in Table 1. By comparing participants' performance in each sentence type, we can find out which type is the most vulnerable to the effects of age and word order.

Sung (2015) focused on the sentence comprehension of elderly adults in Korean, and found significant effects of word order and sentence types. Elderly adults demonstrated greater difficulties in non-canonical sentences than the canonical counterparts. Sung (2015) also found that elderly adults have more difficulties in passive

Table 1
Sentence examples for each condition

CONDITIONS	SENTENCE EXAMPLES
A. Active (2-place)_canonical	Kiiroi hito-ga aoi hito-o osu. yellow person-NOM blue person-ACC push 'The yellow one is pushing the blue one.'
B. Active (2-place)_noncanonical	Aoi hito-o kiiroi hito-ga osu. blue person-ACC yellow person-NOM push 'The yellow one is pushing the blue one.'
C. Active (3-place)_canonical	Aoi hito-ga kuroi hito-ni hako-o ageru. blue person-NOM black person-DAT box-ACC give 'The blue one gives the black one a box.'
D. Active (3-place)_noncanonical	Kuroi hito-ni aoi hito-ga hako-o ageru. black person-DAT blue person-NOM box-ACC give 'The blue one gives the black one a box.'
E. Passive_canonical	Kuroi hito-ga kiiroi hito-ni oikaker-are-ru. black person-NOM yellow person-by chase-PASS-PRS 'The black one is being chased by the yellow one.'
F. Passive_noncanonical	Kiiroi hito-ni kuroi hito-ga oikaker-are-ru. yellow person-by black person-NOM chase-PASS-PRS 'The black one is being chased by the yellow one.'

sentences than active sentences either with two-place or three-place verbs. This result aligns with the previous finding that passive sentences are difficult to process than active sentences, given that passivization involves the movement of the noun phrase with a thematic role from the original, canonical place (Obler et al., 1985, 1991; Rochon et al., 1994; G. S. Waters et al., 1998). These findings suggest that moving a noun phrase with a thematic role from the original structure results in greater processing cost than in active sentences. While Sung (2015) tested only elderly adults, the current study tested Japanese young and elderly adults both to examine the effects of age, word order and sentence types in the same manner as Sung et al. (2017). In sum, our replication study is a combined study of Sung (2015) and Sung et al. (2017); the sentence types used in our study are comparable to Sung (2015) examining both active and passive sentences, but our participants include both young and elderly adults, following Sung et al. (2017).

2. Method

2.1 Participants

A total number of 37 individuals (nine male and 28 female) participated in the study. They were divided into young and elderly groups. The young group included 24 individuals (six male and 18 female). The age ranged from 18 to 24 (mean = 21.2; *SD* = 1.158). The elderly group included 13 individuals (three male and 10 female). Their age ranged from 50 to 75 (mean = 59.7; *SD* = 7.296; 50s: n = 8; 60s: n = 2; 70s: n = 3).[ii] All participants were native speakers of Japanese, and showed a normal range of scores (above 28 out of 30) from the Japanese Mini-Mental State Examination (MMSE) (Mori et al., 1985). This indicates that they do not have any problem with their cognitive function. As for academic background, all participants in the young group were studying in a 4-year university at the time of this study. In the elderly group, one participant completed high school education, nine received a college degree, and other three participants had a graduate degree.

2.2 Materials

2.2.1 Sentence-comprehension task

We employed a sentence-picture matching paradigm in the same way as Sung (2015) and Sung et al. (2017). A participant was presented with picture stimuli that consisted of two pictures, as shown in Figure 1: a picture with the target sentence (e.g., The yellow one is pushing the blue one) and the other picture that displays a reverse thematic role of the actors in the target sentence (e.g., The blue one is pushing the yellow one).

Three types of sentence stimuli were tested: (a) active sentences with two-place verbs, (b) active sentences with three-place verbs, and (c) passive sentences, as shown in Table 1. Following Sung (2015) and Sung et al. (2017), we used three colors of humanized symbols (i.e., "the yellow person," "the blue person," and "the black person") to describe actions in each sentence, which allows all sentences to be semantically reversible.

The stimulus sentences varied by the canonicity of word order. In the canonical condition, NP with a nominative case marker (*ga*) was placed in the sentence-initial position, as in examples A, C, and E in Table 1. In the non-canonical condition, the

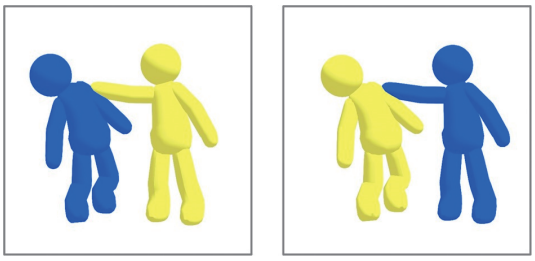

Figure 1
An example of a sentence-comprehension task for the sentence "the yellow one is pushing the blue one" (taken from Sung (2015))

sentence-initial position was filled with the accusative object in a two-place verb sentences, and with the dative object in a three-place verb sentences. In a passive sentence, the agent was placed in this position. See examples B, D and F in Table 1.

The entire task consisted of a total of 36 items with the items belonging to one of six conditions. After the completion of the data collection, it was found that one of the items in the passive-canonical condition had an active verb, resulting in seven two-place verb items in the active-canonical condition and five items in passive-canonical condition.

2.2.2 Procedure

All participants were tested individually in a quiet room. After signing the consent form and filling the demographic questionnaire, MMSE was carried out. Before the main experiment, a practice session was held so that participants familiarize themselves with the components of the experiment. First, participants confirmed whether they understand the meaning of the verbs used in the stimulus sentences. Second, they practiced the sentence-picture matching task. No participant reported difficulties during this training session. In the main experiment session, participants were asked to point to the picture as quickly as possible that corresponds to the auditorily presented stimulus sentences. The experimenter then recorded the response in an answer sheet. When requested, the auditory stimuli were repeated once for the participants. If the first response was incorrect, the stimuli were repeated; this response was independently recorded as the second response.

3. Analysis and Results

Experimental results were recorded for further analysis. The hypothesis in this study was whether age (young vs. elderly), canonicity (canonical vs. non-canonical), voice (active vs. passive), and the number or arguments (two vs. three) have effects on the response pattern. Statistical analyses are performed on data that come from the main experiment and responses obtained from the first response. Although we recorded the second response in case a participant failed to correctly respond in the first trial, following the original studies (Sung, 2015; Sung et al., 2017), we excluded the second response because most of them were correct possibly due to the participants having more time and paying more attention to the task. The experiment did not seek whether participants can produce correct responses but whether they can correctly respond under duress, since we instructed them to respond as quickly as possible. As such, we limit our analyses to the first responses across all the participants.

As in Sung et al. (2017), linear mixed effects model analyses were performed using the *lme4* package and the *lmerTest* package in R. The dependent variable was response (correct or incorrect, coded as 1 or 0). Fixed effects were group by age, canonicity, and voice, and random effects were items and participants. The first model (in Table 2) is based on the full data set. However, a second model (Table 3) is built on a subset of that data (actives only) resulting in those active data points analyzed twice. Further, the model in Table 4 uses another different subset based on age group, and so the data for the elderly group is being used three times. We applied an adjusted alpha-level with a Bonferroni-adjusted p-value. For Table 2 and Table 3, the alpha-level is adjusted to 0.02 (= $0.05/((2+3)/2)$) because the data from the younger group is used twice, and the data from the elderly group is used three times. For Table 4, the alpha-level is adjusted to 0.017 (= $0.05/3$) because the data from the elderly group is used three times.

To find the optimal model, a stepwise procedure

was performed, removing any insignificant fixed effects (largest p-values first) one-by-one (Tredennick et al., 2021). The initial full model tested a three-way interaction between the fixed effects (group* canonicity* voice). Likelihood ratio tests were performed to compare the original model to simpler models with one fixed effect removed each time. When $p > 0.05$, the simpler model was adopted, and the procedure continued. The procedure was performed until $p < 0.05$ in the likelihood ratio tests between the final model and each simpler model. Notably, the final model did not contain the fixed effect of voice. A likelihood ratio test comparing the final model to a model with voice included showed that voice is not significant ($x^2 = 2.22$, $p = 0.14$). This final model is shown in Table 2.

The results in Table 2 show that non-canonical sentences overall received significantly less correct responses ($p < 0.001$). We also find an interaction between age group and canonicity ($p < 0.02$; Bonferroni-adjusted alpha-level). In other words, non-canonical sentences were less correct than canonical sentences in both age groups, but the elderly group had more difficulty with the non-canonical sentences than the younger group. The age group alone does not show significant effect.

In active voice, the experimental stimuli had verbs with two arguments and verbs with three arguments. We tested the interaction between three fixed effects (age group, argument number, and canonicity) and two random effects (item and participant). We followed the same stepwise procedure described above to arrive at an optimal model, which was Table 3.

The intercept shows that the responses were generally correct. The number of arguments alone was not significant in this model ($p = 0.08$). Non-canonical sentences received significantly more incorrect responses ($p < 0.001$).

The performance by the elderly group is reported to be positively affected by education level (Choi, 2012; Sung, 2017). The elderly participants in our study had different educational backgrounds.

Table 2

glmer: response ~ canonicity group + (1 | item) + (1 | participant)*

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.84	0.48	8.00	< 0.001 ***
canonicity = non-canonical	-2.51	0.46	-5.48	< 0.001 ***
group = younger	-0.42	0.54	-0.79	0.43
canonicity = non-canonical:group = younger	1.26	0.51	2.49	0.01 *

Table 3

glmer: response ~ argN group + argN* canonicity + (1 | item) + (1 | participant): this analysis was performed on a subset that only includes active sentences*

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.70	0.56	6.58	< 0.001 ***
arg = N3	-1.05	0.60	-1.75	0.08
group = younger	1.13	0.48	2.35	0.02
canonicity = non-canonical	-2.40	0.53	-4.57	< 0.001 ***
arg = N3: group = younger	-1.07	0.47	-2.25	0.02
arg = N3: canonicity = non-canonical	1.34	0.63	2.13	0.03

Although our sample size is relatively small, we created a subset with data from the elderly group to check whether a suggestion about education level in Sung et al. (2017) is replicated. We tested a model that had two fixed effects (canonicity and education) and one random effect (item). In the results in Table 4, the intercept was elderly participants with a graduate degree. The number of high school graduates was too few to have meaningful judgements. The Boneferroni-adjusted alpha level demonstrates that our data does not show any effect of education level.

In sum, the models demonstrated that non-canonical sentences received significantly more incorrect responses by all age group, and the voice of a sentence did not have effects on the responses. Elderly participants performed with more difficulties when responding to non-canonical sentences. These results from this replication study will be compared with the original study in the next section.

4. Discussion

Recall that Sung (2015) targeted elderly participants only and investigated the effects of sentence types (such as voice and argument numbers) in sentence comprehension; she reported that passive sentences were more difficult than active sentences regardless of the number of verbal arguments. Sung et al. (2017) zoomed in on passive sentences only and they compared results of sentence-comprehension tasks in the elderly group and the younger group. The current replication study is a combined study of Sung (2015) and Sung et al. (2017); our data set is comparable to Sung (2015) examining both active and passive sentences, but our participants include both the elderly group and the young group. In section 4.1, we first discuss the effects of canonicity and age, and show that Sung et al. (2017) and our own study do show some differences when the interaction between the two effects is examined. In section 4.2, we discuss the effects of voice and age, and compare our results with Sung (2015) and Sung et al. (2017) by filtering our data set. In section 4.3, findings that are unique in this replication study are reported.

4.1 Canonicity and age

In both Sung et al. (2017) on Korean and this replication study on Japanese, non-canonical sentences are more difficult than canonical sentences. Sung et al. (2017) investigated the effects of age, canonicity and working memory on passive-sentence comprehension in Korean; we reported our experiment using both active and passive sentences for Japanese young and elderly speakers.

First, the effects of word order in our study were significant in the same way as reported in Sung et al. (2017). Non-canonical sentences were more difficult than canonical sentences; participants regardless of the age group had more incorrect

Table 4

glmer: response ~ canonicity+education+(1 | item): this analysis was performed on a subset that only includes data from elderly adults

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	4.26	0.56	7.60	< 0.001 ***
canonicity = non-canonical	-2.42	0.47	-5.16	< 0.001 ***
education = Highschool	0.00	0.73	0.00	1.00
education = Undergraduate	-0.82	0.40	-2.04	0.04

responses in non-canonical conditions. Sung et al. (2017) argued that canonicity of word order has significant effects on sentence comprehension even in a free word-order language such as Korean. Our results demonstrate that the findings in Sung et al. (2017) extend to the Japanese data. These findings on Korean and Japanese are in line with psycholinguistic studies on the word order effects, which argue that non-canonical word order needs more parsing than canonical word order due to filler-gap dependencies (e.g., Aoshima et al., 2004; Chujo, 1983; Mazuka et al., 2002; Miyamoto & Takahashi, 2004; Tamaoka et al., 2005). Tamaoka et al. (2005) in particular show that it takes longer for young adults to process non-canonical sentences than canonical sentences, even though they can correctly interpret them.

Results in our replication study concerning the interaction between age and canonicity differed from Sung et al. (2017). This original study reported that the interaction between age and canonicity was not significant ($p = 0.400$). In our study, however, a significant interaction between age and canonicity was found ($p < 0.02$; Bonferroni-adjusted alpha level); the elderly group had more difficulty with the non-canonical sentences than the young group. Previous studies on English have shown that sentence-comprehension abilities decline as people get older due to declination of working memory (Caplan et al., 2011; Kemper & Kemtes, 1999; Stine-Morrow et al., 2000), which might also be the reason why Japanese elderly group presented worse performance than the younger group. A comprehensive study on Japanese that includes all these variables (in particular, working memory) regarding sentence comprehension in the elderly group is needed in the future.

4.2 Effects of voice and age

In this section, we discuss how age interacts with voice of a sentence and the number of arguments in a sentence. Sung (2015) tested the data with a repeated measure ANOVA and showed that elderly adults presented worse performance on passive sentences than active sentences either with two-place or three-place verbs in Korean. In our study, we replicated Sung (2015) by filtering the elderly group only, but we differ from Sung (2015) by using a linear mixed effects model with two fixed effects (canonicity and sentence type) and two random effects (item and participant). Our results differ from Sung (2015) in that the elderly group did not show more difficulty by sentence types.[iii] In Table 5, the intercept is active verbs with two arguments and sentences that have canonical word order. The responses in active sentences with a three-place verb ($p = 0.84$) and passive sentences ($p = 0.68$) are not significantly different from the intercept. These results show that the elderly participants in this replication study performed equally regardless of the sentence types.

In Sung et al. (2017), age significantly and negatively correlated with scores on passive-

Table 5

glmer: response ~ canonicity+stype+(1 | item)+(1 | participant); this analysis was performed on a subset that only includes data from elderly adults; Bonferroni-adjusted p-value: 0.017

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.89	0.58	6.73	< 0.001 ***
canonicity = non-canonical	-2.56	0.50	-5.10	< 0.001 ***
stype = Active-3-place-Verb	-0.10	0.48	-0.20	0.84
stype = Passive	0.20	0.50	0.41	0.68

sentence-comprehension task. Since they only investigated the comprehension of passive sentences, we created a subset of our data set with passive sentences only to compare our data with the original study (Table 6). The statistical tests used in Sung et al. (2017) and this replication study were identical, and we found comparable results: canonicity ($p < 0.001$) significantly and negatively correlated with the response scores. In both studies no interaction between age and canonicity was found.

No effects of voice was found in our experiments with Japanese speakers when both voice types (active and passive) were included in the analysis (Table 5). If this is true, we have an interesting situation. Recall that a significant interaction between age and canonicity was found as discussed in the previous section (4.1). If there is no significant interaction between age and voice, it seems to suggest that a syntactic operation that changes surface word order such as scrambling and an operation deriving a passive sentence behave differently with respect to sentence processing. That is, the former creates much greater burden on sentence processing than the latter. It appears then that surface order changing movement plays a more direct role to affect sentence processing than A-movement.

4.3 Effects of number of arguments and education level: new findings

In Sung (2015), the effects of the number of arguments were examined, but the age effects could not be investigated due to the data set that was limited to the elderly group. In Sung et al. (2017), age effects were investigated but the sentence types were limited to passive-sentence comprehension. In our study, this gap is filled by creating a subset of the data set that contained active sentences with two arguments or three arguments (see Table 3). Although we did not find any significant interaction effect between age group and the number of arguments, the model shows that the advantage that young people have in the default case (canonical sentences with two-argument verb; effect size 1.13 in Table 3) disappears in the condition with three arguments. In other words, young people are generally better at processing than older people (in the two-argument condition), but this advantage is neutralized in the three-argument condition: both young and old participants are equally bad.

This finding is intriguing because elderly people have been reported to demonstrate lower sentence-comprehension abilities than young people (e.g., Caplan et al., 2011, 2011; Kemper & Kemtes, 1999; Stine-Morrow et al., 2000). These findings suggest that young people do not necessarily have advantage in comprehending all complex structures. While at this point we do not know why the processing advantage disappears in the three-argument condition, we believe that future studies should not hinge on the assumption that young participants would always perform better than elderly participants.

Table 6

glmer: response ~ canonicity + group + (1 | item) + (1 | participant): this analysis was performed on a subset that only includes passive sentences; Bonferroni-adjusted p-value: 0.02

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.77	0.65	5.83	< 0.001 ***
canonicity = non-canonical	-2.45	0.62	-3.93	< 0.001 ***
group = younger	0.81	0.40	2.02	0.04

5. Conclusion

The present study examined Japanese young and elderly adults' sentence comprehension to replicate Sung et al. (2017) and Sung (2015). Sung et al. (2017) investigated the effects of age, word order and working memory on sentence comprehension in Korean. Sung (2015) studied elderly adults' sentence comprehension alone to explore the effects of sentence types. To the best of our knowledge, our study is the first attempt to examine Japanese elderly adults' comprehension of sentences with scrambled word order. We employed a sentence-picture matching paradigm in the same way as Sung (2015) and Sung et al. (2017) to explore the effects of age, word order, and sentence types on sentence comprehension in Japanese. Our statistical analysis of the experimental results revealed that the effects of age and word order are significant in Japanese in the same way as in Korean. We also found that there was a significant interaction between age and word order. This indicates that elderly adults have more difficulties with non-canonical sentences than young adults. This current replication study of Sung et al. (2017) and to some extent of Sung (2015) not only replicated parts of earlier studies but also contributed to better understanding the nature of the age-related changes in sentence-comprehension abilities in Japanese.

Acknowledgements

This work was partially supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2017S1A2A2038375). The research was also supported by JSPS Kakenhi grants: a Kakenhi (C) #18K00546 to Tomoyuki Yoshida (PI) and Seunghun J. Lee (Co-I), and a Kakenhi (C) #2020K13008 to Tomoe Arii. We thank the participants in the experiment. Data collection was assisted by Atsushi Oho, Takahiro Asayama and Yoon Ah Shin. Any errors in the paper are authors' responsibility.

Notes

(i) Non-canonical sentences are also difficult for Japanese children to comprehend (Hayashibe, 1975; Iwatate, 1980; K. Sano, 1977; Suzuki, 1977). For example, they misinterpret the OSV sentence in (2) as if it were an SOV sentence 'John pushed Mary.' It is not the case, though, that children lack grammatical knowledge of scrambling. When they are provided either with some additional pragmatic or prosodic information, their performance is improved (Minai et al., 2015; Otsu, 1994; T. Sano, 2007).

(ii) Sung et al. (2017) tested 134 participants whose age ranged from 21 to 86 years and were well distributed among various age group. Compared with this previous study, the number of participants of our study is small (n = 37; 50s: n = 8; 60s: n = 2; 70s: n = 3), and our study lacks middle aged participants. As such, this paper divides participants into young and elderly groups, only.

(iii) While the number of elderly participants in Sung (2015) is 51, that of our study is 17. This could have been one of the reasons why results in our study do not align with those from Sung (2015).

References

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language, 51*(1), 23-54. https://doi.org/10.1016/j.jml.2004.03.001

Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging*, *26*(2), 439-450. https://doi.org/10.1037/a0021837

Carpenter, P. A., Miyake, A., & Just, M. A. (1995). Language Comprehension: Sentence and Discourse Processing. *Annual Review of Psychology*, *46*(1), 91-120. https://doi.org/10.1146/annurev.ps.46.020195.000515

Choi, H. (2012). Expressive language function in the normal elderly: Effects of gender and educational level. *Special Education Research*, *11*(1), 131-146.

Chujo K. (1983). Nihongo tanbun-no rikai katei — Bunrikai sutoratejii no sougo kankei [The

Interrelationships among strategies for sentence comprehension]. *The Japanese Journal of Psychology*, *54*(4), 250-256. https://doi.org/10.4992/jjpsy.54.250

DeDe, G., Caplan, D., Kemtes, K., & Waters, G. (2004). The Relationship between age, verbal working memory, and language comprehension. *Psychology and Aging*, *19*(4), 601-616. https://doi.org/10.1037/0882-7974.19.4.601

Fodor, J. D., & Hirose, Y. (2003). What Japanese parsing tells us about parsing. In W. McClure (Ed.), *Japanese Korean Linguistics 12* (pp. 193-295). CSLI Publications.

Fodor, J. D., & Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research*, *23*(5), 407-434. https://doi.org/10.1007/BF02143947

Fodor, J. D., & Inoue, A. (1998). Attach anyway. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 101-141). Kluwer.

Hayashibe, H. (1975). Word order and particles: A developmental study in Japanese. *Descriptive and Applied Linguistics*, *8*, 1-18.

Inoue, A., & Fodor, J. D. (1995). Information-paced parsing of Japanese. In R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing* (pp. 9-63). Lawrence Erlbaum.

Iwatate, S. (1980). The word-order and case strategies in Japanese children. *The Japanese Journal of Psychology*, *51*, 233-240.

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

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133-156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kamide, Y., & Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes*, *14*(5-6), 631-662. https://doi.org/10.1080/016909699386211

Kemper, S., & Kemtes, K. (1999). Limitations on syntactic processing. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: Aging, grammar, and memory* (pp. 79-105). Kluwer Academic.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676-703. https://doi.org/10.1037/0033-295X.101.4.676

Mazuka, R., & Itoh, K. (1995). Can Japanese speakers be lead down the garden path? In R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing* (pp. 295-329). Lawrence Erlbaum.

Mazuka, R., Itoh, K., & Kondo, T. (2002). Cost of scrambling in Japanese sentence processing. In M. Nakayama (Ed.), *Sentence processing in East Asian languages* (pp. 131-166). CSLI Publications.

Minai, U., Isobe, M., & Okabe, R. (2015). Acquisition and use of linguistic knowledge: Scrambling in child Japanese as a test case. *Journal of Psycholinguistic Research*, *44*(3), 287-307. https://doi.org/10.1007/s10936-014-9347-x

Miyamoto, E. T., & Takahashi, S. (2004). Filler-gap dependencies in the processing of scrambling in Japanese. *Language and Linguistics*, *5*(1), 153-166.

Mori, E., Mitani, Y., & Yamadori, A. (1985). Shinkeishikkankanja niokeru nihongo ban Mini-Mental State tesuto no yukosei [The validity of Japanese Mini-Mental State Examination for patients suffering neurological disorder]. *Shinkei Shinri-gaku [Japanese Journal of Neuropsychology]*, *1*, 82-90.

Morris, R. G., Gick, M. L., & Craik, F. I. M. (1988). Processing resources and age differences in working memory. *Memory & Cognition*, *16*(4), 362-366. https://doi.org/10.3758/BF03197047

Nakatani, K., & Gibson, E. (2009). An on-line study of Japanese nesting complexity. *Cognitive Science*, *34*(1), 94-112. https://doi.org/10.1111/j.1551-6709.2009.01067.x

Nakayama, M. (1995). Scrambling and probe recognition. In R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing* (pp. 257-273). Erlbaum.

Nemoto, N. (2017). Scrambling. In N. Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 121-153). Blackwell Publishing Ltd. https://doi.org/10.1002/9781405166225.ch5

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics*, *12*(4), 433-452. https://doi.org/10.1017/S0142716400005865

Obler, L. K., Nicholas, M., Albert, M. L., & Woodward, S. (1985). On Comprehension Across the Adult Lifespan. *Cortex*, *21*(2), 273-280. https://doi.org/10.1016/S0010-9452(85)80032-0

Otsu, Y. (1994). Early acquisition of scrambling in Japanese. In T. Hoekstra & B. D. Schwartz (Eds.), *Language acquisition and language disorders* (Vol. 8, pp. 253-264). John Benjamins Publishing Company. https://doi.org/10.1075/lald.8.12ots

Rochon, E., Waters, G. S., & Caplan, D. (1994). Sentence comprehension in patients with Alzheimer's disease. *Brain and Language*, *46*(2), 329-349. https://doi.org/10.1006/brln.1994.1018

Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications.* [Doctoral dissertation]. Massachusetts Institute of Technology.

Sano, K. (1977). An experimental study on the acquisition of Japanese simple sentences and cleft sentences. *Descriptive and Applied Linguistics*, *10*, 213-233.

Sano, T. (2007). Early acquisition of copy & movement in a Japanese OSV sentence. *BUCLD 31 Online Proceedings Supplement*. http://www.bu.edu/bucld/ proceedings/supplement/vol31/

Stine-Morrow, E. A. L., Ryan, S., & Leonard, J. S. (2000). Age differences in on-line syntactic processing. *Experimental Aging Research*, *26*(4), 315-322. https://doi.org/10.1080/036107300750015714

Sung, J. E. (2015). Effects of syntactic structure on sentence comprehension abilities as a function of the canonicity of word-order and their relation to working memory capacity in Korean-speaking elderly adults. *Communication Sciences & Disorders*, *20*, 24-33.

Sung, J. E. (2017). Age-related decline in case-marker processing and its relation to working memory capacity. *The Journals of Gerontology: Series B*, *72*(5), 813-820. https://doi.org/10.1093/geronb/gbv117

Sung, J. E., Yoo, J. K., Lee, S. E., & Eom, B. (2017). Effects of age, working memory, and word order on passive-sentence comprehension: Evidence from a verb-final language. *International Psychogeriatrics*, *29*(6), 939-948. https://doi.org/10.1017/S1041610217000047

Suzuki, S. (1977). Nihon no yoji ni okeru gojun horyaku [Word-order strategy by Japanese children]. *The Japanese Journal of Educational Psychology*, *25*(3), 56-61.

Tamaoka, K., Sakai, H., Kawahara, J., Miyaoka, Y., Lim, H., & Koizumi, M. (2005). Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research*, *34*(3), 281-332. https://doi.org/10.1007/s10936-005-3641-6

Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, *102*(6). https://doi.org/10.1002/ecy.3336

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155-179). Lawrence Erlbaum.

Waters, G., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, *16*(1), 128-144. https://doi.org/10.1037/0882-7974.16.1.128

Waters, G., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 550-564. https://doi.org/10.3758/BF03195534

Waters, G., & Caplan, D. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory*, *13*(3-4), 403-413. https://doi.org/10.1080/09658210344000459

Waters, G., Rochon, E., & Caplan, D. (1998). Task demands and sentence comprehension in patients with dementia of the Alzheimer's type. *Brain and Language*, *62*(3), 361-397. https://doi.org/10.1006/brln.1997.1880

Yeom, T. H., Park, Y. S., Oh, K. J., Kim, J. K., & Lee, Y. H. (1992). *A manual for K-WAIS*. Korean Guidance.

Zurif, E., Swinney, D., Prather, P., Wingfield, A., & Brownell, H. (1995). The allocation of memory resources during sentence comprehension: Evidence from the elderly. *Journal of Psycholinguistic Research*, *24*(3), 165-182. https://doi.org/10.1007/BF02145354