Investigation of the Relationship Between Animacy and L2

Learners' Acquisition of the English Plural Morpheme

Abstract

Languages have been argued to differ in the extent to which they allow plural forms of nouns according to the Animacy Hierarchy. Japanese distinguishes between animate and inanimate nouns; the latter is less likely to receive plural markers (e.g., ?*hon-tachi*), unlike English. This L1-L2 difference might cause difficulty in acquiring the plural morpheme. The present study thus investigates the influence of animacy on the processing of the English plural morpheme in online sentence processing. In this study, 34 Japanese university students engaged in a moving window-version of the self-paced reading task, during which they also judged whether the number of words presented was one or two when prompted. If animacy matters, Japanese EFL learners might not show an interference effect of this second task for inanimate nouns. However, as no such effect was found, the prediction based on the Animacy Hierarchy is not borne out.

Keywords: Animacy Hierarchy, L1 influence, Stroop-like number judgment task, plurality, number information

Background

Research on the L2 Acquisition of the Plural Morpheme

Since the beginning of research in the field of second language acquisition, the issue of the order in which second language (L2) learners acquire grammatical morphemes (e.g., progressive be+*-ing*, possessives, third-person singular *-s*, plural *-s*, past tense *-ed*) has received considerable attention (Bailey et al., 1974; Dulay & Burt, 1974; Larsen-Freeman, 1975; Hakuta, 1976). Among several factors that might have an impact on the acquisition of morphemes, learners' first language (L1) has been the subject of numerous previous studies (e.g., Luk & Shirai, 2009; Murakami & Alexopoulou, 2016).

Researchers have considered the plural morpheme to be acquired comparatively early in L2 development (Murakami & Alexopoulou, 2016). However, in sentence processing research, L2 learners are known to show insensitivity to the presence of the plural morpheme (Jiang, 2004, 2007). For example, Jiang (2004) conducted a self-paced reading study focusing on the agreement attraction paradigm. The participants in his study read sentences like *the key to the cabinet(s) <u>was</u> rusty*. Native speaker participants read significantly more slowly when the noun in the prepositional phrase was plural than when it was singular, whereas L2 participants did not show such a tendency. On the basis of this and similar experimental results, Jiang attributed the discrepancies between native and L2 learners to the lack of grammatical knowledge of L2 learners.

Arguably, the difficulty of acquiring certain grammatical items is determined by whether the learner's L1 has a language system similar to that of the target language. In particular, there are many reports of learners who seem to behave similarly to native speakers yet show differences in their immediate reactions to grammatical errors (e.g., Jiang, 2004, 2007). The key issue here is that L2 learners know the rule that when

multiple entities are referred to, plural morphology should be attached to the noun. However, when they are engaged in a meaning-focused comprehension task, they demonstrate insensitivity to the plural morpheme (Jiang, 2004; Jiang et al., 2017). This kind of performance difference is also observed in sentence-processing studies focusing on the acquisition of the English third-person singular morpheme -*s* (Bannai, 2011; Shibuya & Wakabayashi, 2008).

The cause of this difficulty is considered to stem from the learners' L1, which led Jiang et al. (2011) to propose the Morphological Congruency Hypothesis (MCH). The MCH posits that if learners' L1 does not have a corresponding morpheme, it is difficult for them to acquire a given morpheme in their L2. Jiang et al. (2011) conducted self-paced reading experiments similar to Jiang (2007) and compared Russian L2 learners of English and Japanese L2 learners of English. The results demonstrated that while Russian learners, whose L1 has plural morphology, showed sensitivity to plurality errors, Japanese learners did not, on which basis Jiang et al. (2011) proposed the MCH. However, several previous studies reported evidence against the MCH and argued that agreement conditions (Song, 2015), L2 proficiency (Wen et al., 2010), and/or mass/count distinctions (Choi & Ionin, 2021) play an important role in the acquisition of plural morphemes.

As for the acquisition of plural morphology, it has been observed that English learners whose L1 is Japanese or Chinese, where plurality is not obligatorily marked, have greater difficulty in acquiring plural morphemes than do learners whose L1 is Russian, where plurality is obligatorily marked (Jiang et al., 2011). Jiang et al. (2011) concluded that the difference between Russian and Japanese learners of English lies in whether the learners' L1s allow number marking on nouns. However, the cause could be

attributed not just to number marking on nouns but to the kind of nouns that can be pluralized, which can be addressed within the framework of the Animacy Hierarchy (Corbett, 2000; Haspelmath, 2013). This will be reviewed in the next section.

The Animacy Hierarchy

In the framework of the Animacy Hierarchy of nouns, including pronouns, proposed by Corbett (2000), there is a difference among languages in the acceptability of noun plurality depending on the animacy of the noun; the boundaries of this hierarchy differ between languages, as shown below.[1]

1st person > 2nd person > 3rd person > kin > human > animate > inanimate

For example, Japanese allows nouns to receive plural markers for non-human animate objects (e.g., *inu-tachi* 'dogs', *tori-tachi* 'birds'). According to the Animacy Hierarchy, if a language allows the pluralization of a lower order of the hierarchy, it allows the pluralization of all the higher-order classes. Thus, in the case of Japanese, 1st person, 2nd person, 3rd person, kin, and all other human nouns may be pluralized, though pluralization of nouns is not obligatory in Japanese.

A significant difference between pluralization in English and Japanese is that the plural markers in Japanese are not grammatical morphemes but suffixes attached to nouns since there is no grammatical configuration such as number agreement involving inflectional morphemes in Japanese (Kato, 2006). Notably, although Corbett (2000) proposed that Japanese does not allow pluralization of inanimate nouns, recent research

---

[1] There has been a recent attempt to identify a hierarchical ordering of uncategorized inanimate nouns based on the similarities to the characteristics of human beings. See Ji and Liang (2018) for a detailed discussion of animacy.

on the use of the Japanese plural marker -*tachi* has shown that inanimate nouns can be

pluralized (Murahata, 2019). In addition, another Japanese plural marker -*ra* can be

attached to inanimate nouns, though very occasionally (e.g., *sakuhin-ra* 'work-PL').[2]

Nevertheless, -*ra* can be attached to demonstratives such as *kore-ra* 'these' or *sore-ra*

'those', whereas -*tachi* cannot (e.g., **kore-tachi*; see Morita, 1980, pp. 267–268, for other

differences of these Japanese plural markers).

In contrast to English, some languages do not allow the pluralization of animate

nouns, an example of which is Chinese. Chinese has the pluralization marker -*men*,

although it is debatable whether Chinese -*men* is a plural marker or collective marker, the

latter meaning 'one representative and some others' (Iljic, 1994; Li, 1999). Similarly,

Japanese -*tachi* can have a collective interpretation, such that *Taro-tachi* does not mean

two or more people named *Taro*, but rather a group of people, the representative of whom

is in some way *Taro* (Nakanishi & Tomioka, 2004). However, unlike Japanese -*tachi*,

-*men* can only be attached to pronouns and human nouns but not to other animate nouns,

such as those for animals. Another difference between Chinese -*men* and Japanese -*tachi*

lies in definiteness (Iljic, 1994; Li, 1999). Despite Japanese -*tachi* not ruling out the

indefinite interpretation in some cases (Nakanishi & Tomioka, 2004), adding -*men* to a

noun in Chinese always makes it definite. Therefore, Lardiere (2009, 2010) argued that

Chinese learners of English need to disassociate definiteness and the plural marker and

"reassemble" them onto the definite article and plural morpheme in English, which she

calls *feature reassembly*.

---

[2] Consistent with Murahata (2019), examples of pluralized inanimate nouns were found in NINJAL Web Japanese Corpus (NWJC; Asahara et al., 2016). However, there are few such cases, and most of the nouns pluralized by -*tachi* and -*ra* are animate nouns, particularly human nouns. See the Supplementary Material for further details.

One of the similarities between the Chinese and Japanese number representation systems is that the absence of the plural marker does not necessarily mean that the noun has a singular meaning, which Corbett (2000) called *general number* in these two languages. Languages that have a general number system do not refer to the number of nouns in expressing their meaning. For example, a Japanese noun without *-tachi* means that its number is 'one or more' or 'just one'. English, conversely, does not have a general number and must express either the singular or the plural.

In contrast to Japanese and Chinese, some languages, such as English and Russian, allow the pluralization of inanimate nouns. Russian is similar to English in that both have morphologically plural forms and allow inanimate nouns to be morphologically plural. Nevertheless, one difference between these two languages is that in English, the singular form is always the base (stem) form, and the plural form is created by adding inflectional morphemes. Contrarily, Russian's singular forms do not correspond to the stem and have both singular and plural inflections (e.g., *komnat-a* is the singular form of 'room' and *komnat-y* is its plural form). In addition, some nouns that can be pluralized in English, such as *potato* and *grape*, do not have plural forms in Russian (e.g., *kartofel'* 'potato', *vinograd* 'grape'; Corbett, 2012).

To summarize, according to the Animacy Hierarchy, the difficulty of learning plurality is determined by the difference in noun categories. For Japanese learners of English, it is predicted that they would be sensitive to the difference between singular and plural forms of animate nouns but not to that of inanimate nouns since Japanese does not on a regular basis exercise pluralization of inanimate nouns.

Motivation for the Study

Although L2 acquisition of the mass/count distinction has been investigated, there has been little discussion of the effects of other features of nouns on the acquisition of the plural morpheme. Among them, one important feature that should be considered is animacy, which plays an important role in the processing of language (e.g., Lempert, 2016), and the way nouns express plurality varies greatly across languages (Corbett, 2000). Particularly relevant to the discussion of the present study is the Animacy Hierarchy (Corbett, 2000; Haspelmath, 2013), according to which languages can be divided based on which types of nouns in the hierarchy are differentiated in number. In English, for example, the lowest in the hierarchy (inanimate nouns) can be pluralized, except for mass nouns (e.g., *sand*, *water*). In contrast, the Japanese optional plural marker *-tachi* can be attached to animate nouns, while the pluralization of inanimate nouns is very rare (?*tsukue-tachi* 'desk -PL'). Recall that the MCH postulates that the acquisition of L2 morphemes is based on whether the learners' L1 has the same morpheme. However, Jiang's (2004) seminal study reporting L2 learners' insensitivity to the English plural morpheme has been criticized for the fact that most of the nouns used in the test sentences were inanimate nouns (Lempert, 2016). In fact, as Haspelmath (2013) stated, the pluralization of nouns varies across languages not only in terms of the obligatoriness of pluralization but also animacy. In other words, if the insensitivity to the plural morpheme in the previous studies depended on the animacy of nouns, consideration of the Animacy Hierarchy will ensure a better and more precise prediction regarding the acquisition of the plural morpheme. Contrarily, if the responses to plural nouns do not show differences based on the animacy of the nouns, it can be concluded that the insensitivity to the plural morphemes could be due to other factors related to the pluralization of nouns, such as the mass/count distinction, that are considered difficult for

L2 learners to acquire (Choi et al., 2018; Choi & Ionin, 2021; Inagaki, 2014; Tsang, 2017).

Another limitation of previous research is methodological: To date, research on the acquisition of pluralization of nouns has tended to focus on whether the learners detect either specifier-head number agreement (e.g., *many boat/s a lot of new computer/s*) or subject-verb number agreement (e.g., *the officer/s from the station/s is/are*; Choi & Ionin, 2021; Choi et al., 2018; Jiang, 2004, 2007; Lempert, 2016), except for Jiang et al. (2017), who used a sentence-picture matching task. However, the mechanism of L2 learners' grammatical knowledge and their processing tendency should also be investigated by examining how grammatical (well-formed) sentences are processed (Trenkic et al., 2014). Moreover, as pointed out by Author (2018), as computing number agreement involves more than just the knowledge of plurality of nouns, it is difficult to pinpoint the locus of learners' insensitivity to plural morphemes. The insensitivity to the plural morpheme found in Jiang (2004) might reflect either that (a) the L2 learners did not process the plural morpheme or that (b) they processed the plural morpheme but failed to integrate it with syntactic number agreement. The former case could be investigated by using a lexical decision task and comparing the reaction times for singular and plural nouns. If the plural nouns are processed more slowly, it implies that the learners decompose the plural morpheme, which requires additional processing time. Nevertheless, demonstrating that the learners process the plural morpheme does not guarantee that they are able to compute number agreement. For successful number agreement, learners need to associate the decomposed plural morpheme and plurality. Thus, in this study, "acquisition" is defined as making form-meaning mappings such that the English plural morpheme *-s* is associated with its meaning of plurality, or 'more than one'. To investigate the

form-meaning mapping, the present study adopts a novel psycholinguistic experiment called the Stroop-like number judgement task, originally proposed by Berent et al. (2005). In this task, the participants are required to judge whether the number of words presented on a screen is one or two. In the target condition, the number feature of nouns and the presented number of words mismatch such that a noun is presented in a plural form, and the participants need to judge it as one word. Berent et al. (2005) found that the participants' response was significantly slower when judging plural nouns than singular nouns as one word. This slower response time (RT) implies the interference of noun plurality in judging the number of words. Later, Patson and Warren (2010) expanded this Stroop-like number judgment task to a reading task in which the participants read a sentence presented by either one word or two word chunk and were asked to judge the number of words when prompted. Patson and Warren (2010) succeeded in replicating the results of Berent et al. (2005) and found that the paradigm can be expanded from single-word recognition experiments to reading tasks.

The advantage of this experimental paradigm over the anomaly detection paradigm in which participants' responses to grammatical errors are investigated is that the task enables us to directly measure whether the participants understand the meaning of plural nouns. Moreover, the Stroop-like nature of the task also helps overcome the distinction between offline and online measures. It is known that L2 learners show performance discrepancies in online (e.g., self-paced reading task, eye-tracking) and offline tasks (e.g., grammaticality judgment task, acceptability judgment task; e.g., Ellis, 2005, 2006; Suzuki, 2017; Vafaee et al., 2017). The nature of the online measures is meaning-based comprehension. The participants are instructed to focus on comprehending the meaning of the sentence rather than focusing on whether the sentence is grammatical. Therefore,

the online measures are considered to inhibit access to conscious/explicit knowledge. Since the Stroop task does not directly ask the learners to judge the grammaticality of the sentence or direct their attention to form, it is believed that the RT delay found in judging plural nouns as one word reflects the implicit and automatic processing of the plural morpheme.

Using the Stroop-like number judgment task, the purpose of the present study is to investigate Japanese learners' acquisition of the English plural morpheme by manipulating the animacy of nouns.

The Present Study

Research Questions and Predictions

The research question of the present study is:

Does Japanese L2 learners' acquisition of the English plural morpheme depend on the animacy of the noun?

If the L1 influence is present, as predicted by the MCH, and the Animacy Hierarchy influences the acquisition of the plural morpheme, Japanese L2 learners will be sensitive to the plurality only of animate nouns, not of inanimate nouns.

To investigate the acquisition of the plural morpheme, the present study adopts the Stroop-like number judgement task proposed by Berent et al. (2005). The original Stroop-like number judgment task involves presenting either a single word or multiple words on the screen and asking the participants whether one or two words are presented. However, this mode of presentation has a potential risk: The participants might not

process the language but strategically focus on visual information such as the space between the two words in an attempt to judge the number of words as quickly as possible. Therefore, for the learners to engage in processing the linguistic information and maintain the meaning-focused processing that is critical in assessing the implicit nature of the learners' knowledge, as reviewed earlier, the present study used the sentence reading version of the Stroop-like number judgment task. In this task, the participants were first engaged in word-by-word version of the self-paced reading task. During the task, when the target word appeared on the screen, the screen color changed, which is the prompt for the participants to judge whether the presented word or words is/are either one word or two words.[3] The validity of this task in investigating L2 learners' acquisition of plural morphemes was confirmed by Author (2018), which found a significant RT delay in judging plural nouns compared to singular nouns for both native speakers of English and Japanese learners of English (see Section 3.4 for a detailed explanation of the task).

Participants

The participants of this study were 34 Japanese graduate or undergraduate students, with a mean age of 19.67 ($SD = 1.98$; $Min = 18$, $Max = 25$), of whom 29 had had experience studying abroad, with a range of a quarter to 12 months ($M = 3.17$, $SD = 4.23$, $Mdn = 1$). Based on the average score of the Oxford Quick Placement Test (OQPT; $M = 37.88$, $SD = 6.38$ out of a full score of 60), the proficiency level of most of the participants was B1 on the scale of the Common European Framework of Reference (though the range is from A2 to C1). The reliability coefficient of the OQPT was 0.75 [0.63, 0.87].

---

[3] For those who are not familiar with this task, please watch the video uploaded on the following: https://osf.io/chnrj/?view_only=012ea92af42f41679bd6011f54ef00e4

Materials

Most of the test sentences were borrowed from Author (2018) and revised to control the animacy of nouns. All the nouns were regularly inflected nouns that make plural forms by attaching -*s*, and no irregularly inflected nouns (e.g., *man-men*, *potato-potatoes*, *strawberry-strawberries*) were included as target nouns. Since it was difficult to control the frequency level of the target nouns, particularly due to the limited number of highly frequent animate nouns, a measure of noun frequency was added as a covariate in analyzing the data (see section 3.4). There were 40 sets of test sentences, an example of which is shown in (1a–d).

(1a) Every boy in school admired the excellent <u>bikes</u>.

(1b) Every boy in school admired the excellent <u>bike</u>.

(1c) Every boy in school admired the excellent <u>players</u>.

(1d) Every boy in school admired the excellent <u>player</u>.

All the target nouns were placed at the end of the sentence in the same structure, *the* + Adj + N. The target nouns were inanimate in (1ab), while animate nouns were used in (1cd). Across the four conditions, (1ac) include plural nouns, which are expected to induce slower responses than singular nouns in (1bd) due to the interference of plurality in the Stroop-like number judgment task. The test sentences were distributed into four

lists so that each sentence was presented in only one of the four conditions. Specifically, the participants did not see the same sentences more than once.

If the animacy of nouns matters in the acquisition of the English plural morphemes, it is predicted that processing the plural morpheme attached to inanimate nouns (e.g., *bikes*) would be more difficult than that attached to animate nouns (e.g., *players*). Therefore, a significant RT difference would be observed only between (1c) and (1d) but not between (1a) and (1b).

In addition to 40 test sentences, there were 80 filler items, in 60 of which the target nouns were presented as two words, and the correct answer for the rest was "one." In addition, the position of the number judgment was set in the middle of the filler sentences to prevent the participants from guessing when the number judgment would be required. In total, the participants were required to judge 120 sentences, half of which required "one" as a correct answer and the other half "two."

After responding to all 40 target items and 40 of the 80 fillers, the participants were asked to answer a true-or-false question about the sentence that they had just read. Half of the questions required the "True" response, while the other half required "False."

Tasks and Procedures

The Stroop-like number judgment task was developed by the author using Hot Soup Processor version 3.4 (https://hsp.tv/). In this task, the participants were required to perform a moving-window version of the self-paced reading task. To proceed to the next stage, the participants were required to press the space bar. Two hundred milliseconds after the target word(s) was presented on the screen, the screen color changed to blue, which signaled that the participants were required to judge whether the words presented

on the screen numbered one or two by pressing arrow keys as quickly as they could. The measurement of the reaction time started when the screen color changed to blue, which was 200 ms after the target word was presented, and stopped when the participant pressed either of the arrow keys. After reading the sentence, the participants were required to answer a simple true or false question about the sentence they had just read. The participants were allowed to move on to the next item at their own pace. **Error! Reference source not found.** visually represents the reading part and the number judgment part.

[Insert Figure 1 around here]

The experiment was conducted individually. Before the experiment began, the participants were asked to give informed consent and agreed to participate in the session with a compensation of 2000JPY. First, the participants took a paper-based Oxford Quick Placement Test (OQPT), which consists of 60 items, within 30 minutes. After a short break, the participants completed the Stroop-like number judgment task, which lasted approximately 20 to 25 minutes. Subsequently, the participants performed two other experimental tasks irrelevant to the present study. At the end of the session, the participants answered a background questionnaire. The entire session took approximately two hours.

Analysis

The RT data were analyzed as follows. First, the accuracy of comprehension questions was calculated for both item and subject, showing that the accuracy of the

16

comprehension questions for Item No. 2 was lower than 50%. A closer look at this item revealed an error in the comprehension questions. Thus, all the responses to this item were excluded. The mean accuracy of the comprehension questions after removing Item No. 2 was high ($M = 89.7\%$, $SD = 7.18\%$), suggesting that the participants focused on meaning comprehension. However, examining the individual data revealed that the accuracy of one of the participants on the comprehension questions was 71.8%, which might indicate that that participant did not focus on meaning-based comprehension during the task. Therefore, this participant was removed from further analysis.

Second, all the responses in which the participants failed to answer comprehension questions were removed from the data set. The removed responses accounted for 9.8% of the data. Subsequently, the overall accuracy of number judgment was investigated, and the incorrect number judgment responses were removed for the RT analysis. Table 1 summarizes the error rate of the number judgment across the four conditions.

[Insert Table 1 around here]

Third, responses longer than the individual mean RT±3SD or longer than 2000ms were treated as outliers and removed. The 2000ms cut-off point was determined on the basis of the visual inspection of the RT distribution (see the Supplementary Material). The removed responses accounted for 3.7% of all the correct number judgments.

After these data cleaning procedures, the data were submitted to generalized linear mixed-effects model using R version 4.1.2 (R Core Team, 2021) and *lme4* package version 1.1-27.1 (Bates et al., 2021). The predictor variable was the raw RT of the number judgment data. Furthermore, the explanatory variables were the number feature

of nouns (singular or plural) and animacy of nouns (animate or inanimate). Both categorical variables were re-coded by sum-contrast coding (number: −0.5 = singular, 0.5 = plural; animacy: −0.5 = animate, 0.5 = inanimate). In addition to these two variables of interest, the following six covariates that might have an impact on the RT were also taken into account: the number of letters, the number of syllables, frequency of singular forms, frequency of plural forms, cumulative frequency (the sum of the frequency of singular and plural forms), and the presentation order.

The frequency information was extracted from the SUBTLEX-US corpus (Brysbaert & New, 2009) and was transformed to the Zipf scale, which is "log10 (frequency per million words) + 3" (van Heuven et al., 2014, p. 1179) for a better fit (see the Supplementary Material for the frequency information of the target nouns). Notably, all these numerical covariates, including word length, syllables, and presentation order, were z-transformed using the grand means before entering the model, to avoid convergence issues.

Model building was conducted as follows. First, null models that only included by-subject and by-item random intercepts were fit using either the gamma distribution or the inverse Gaussian distribution (with identity link function). Both the intercepts are considered to fit raw RT data (Lo & Andrews, 2015). Model comparison based on the Akaike Information Criterion (AIC) demonstrated that the model with the inverse Gaussian distribution better fit the current data. Thus, the following analysis was conducted using a combination of the inverse Gaussian distribution and identity link function.

Second, before examining the effects of number marking and animacy, the main effects of the covariates described above were examined in a forward manner (see the

Supplementary Material for details). As a result, adding presentation order, the frequency of singular forms, and the word length (the number of letters) improved the model fit the best; thus, these three covariates were added to the model.

The best fit random effect structure was then determined using backward elimination. First, the model with a maximal random effect structure was built, and then simplified according to the results of the principal component analysis of the random effect structure, using the *reCPA* function in the *lme4* package. The least explained random effect component was removed, and the model was refit until it was confirmed that all the random effect components contributed to the model. The final model included the by-item random slope of animacy, noun number, and the interaction between the two, the by-subject random slope of noun number, and the interaction term besides the by-subject random intercept. To investigate the interaction between animacy and the noun number, a simple main effect test was conducted using the *emmeans* function in the *emmeans* package version 1.7.0 (Lenth et al., 2021). All the R code and its output, including figures not included in this paper, are available at Open Science Framework (https://osf.io/chnrj/?view_only=012ea92af42f41679bd6011f54ef00e4).

Results

Table 2 summarizes the descriptive statistics of the RT across the four experimental conditions. It appears that for both animate and inanimate nouns, the participants took longer to respond to the target plural nouns than the singular nouns.

[Insert Table 2 around here]

The results of final GLMM model demonstrated that while the Zipf cumulative frequency is significant (*Estimate* = −15.22, *SE* = 6.08, *t* = −2.50, *p* = .012), the other two covariates were not (presentation order: *Estimate* = −2.60, *SE* = 5.42, *t* = −0.48, *p* = .631; word length: *Estimate* = −0.69, *SE* = 6.26, *t* = −0.11, *p* = .912). The main effect of the cumulative frequency indicates that the participants responded more quickly to the highly frequent nouns. Interestingly, the main effect of number marking was significant (*Estimate* = −65.63, *SE* = 19.80, *t* = 3.32, *p* = .001), while the main effect of animacy was not (*Estimate* = 6.26, *SE* = 13.47, *t* = 0.47, *p* = .642). However, the interaction of the number marking and animacy did not reach statistical significance (*Estimate* = 42.96, *SE* = 33.52, *t* = 1.28, *p* = .200). Table 3 summarizes the results of the final GLMM model, and Figure 2 visually represents the predicted RTs in each condition. It should be noted that the RTs presented in Figure 2 are different from the mean RTs presented in Table 3 because Figure 2 is based on the model estimate and does not describe the original data collected in the experiment.

Although the interaction term was not significant, the planned simple main effect test was conducted and revealed that for inanimate nouns, plural nouns were judged more slowly than singular nouns (*Estimate* = −87.1, *SE* = 26.5, *z* = −3.28, *p* = .001). Conversely, for animate nouns, the difference was smaller and did not reach significance (*Estimate* = −44.2, *SE* = 25.3, *z* = −1.74, *p* = .081). However, care should be taken in interpreting the difference between animate and inanimate nouns, as the *p*-value is close to .05.

[Insert Table 3 around here]

[Insert Figure 2 around here]

Collectively, these results suggest that the processing of plural information did not differ by the animacy of the noun. However, the singular-plural difference may be influenced by the animacy of nouns. These results contradict the prediction based on the Animacy Hierarchy, as will be further discussed below.

Discussion

The purpose of the present study was to investigate the influence of animacy on Japanese learners' acquisition of English plural morphemes. It was hypothesized that Japanese learners of English would be sensitive to the singular/plural distinction of animate nouns but not inanimate nouns. This is because Japanese does not allow pluralization of inanimate nouns. Surprisingly, however, the results of the present study indicated that the participant's responses to the number judgment of plural nouns were significantly slower than for singular nouns, irrespective of the animacy of nouns.

The results of the experiment demonstrate that L2 learners whose L1 has a different mechanism of pluralization of nouns described by the Animacy Hierarchy were sensitive to the addition of the plural morpheme, regardless of whether nouns are animate or inanimate. Furthermore, this finding is contrary to the MCH (Jiang et al., 2011) that suggested the insensitivity to the plural morphemes for Japanese L2 learners of English. There are several possible explanations for this result. First, it is possible that the learners who showed sensitivity to the plural morpheme in the number judgement task would not respond to number agreement errors in online sentence processing. In particular, the discrepancies between the current study and Jiang et al. (2011) could be due to a task effect. As mentioned in the literature review, previous studies of the acquisition of the

plural morpheme mostly focused on anomaly detection of number agreement errors in sentence processing (Jiang, 2004, 2007; Jiang et al., 2011; Song, 2015; Wen et al., 2010). In this study, however, the task did not measure the learners' sensitivity to number agreement errors, given that number agreement involves more complex grammatical processing than the processing of plural morpheme. Thus, bridging the gap between the acquisition and processing of the plural morpheme and the use of that knowledge in number agreement is an important issue for future research.

Similarly, another possible explanation for the result is the degree of attention to form during the task engagement. Although the participants were not instructed to pay attention to the form during the task, such as singular-plural differences of nouns, the Stroop effect (mismatch in number) occurring during the process of number judgment could have forced the participants to focus more on the number features of nouns. Nevertheless, if the learners had been aware that the number feature is the key to the task, they could have ignored the number feature of nouns, such as whether the plural morpheme is attached or not. The fact that the learners' judgement was still influenced by the plurality of nouns implies that the RT delay resulted from the uncontrollable activation of the plural meaning. In other words, the learners could not ignore the number feature because its activation was automatic, happening within a second on average.

Another interesting finding is the lack of a difference between animate and inanimate nouns. Based on the Animacy Hierarchy proposed by Corbett (2000), it was predicted that the RT delay would be observed only for animate nouns but not for inanimate nouns, given that in Japanese, the L1 of the participants in the current study, the pluralization of inanimate nouns is very rare. Nevertheless, the results indicate that the Japanese learners of English were sensitive to the plural morpheme for both animate

and inanimate nouns. In other words, the impact of animacy was less critical in the current experimental task than was expected in other sentence processing research (e.g., Lempert, 2016).

It should, however, be noted that the observed singular-plural RT difference was larger when participants processed inanimate nouns, which seems to constitute a "reverse" animacy effect because the rare pluralization of Japanese inanimate nouns was expected to cause difficulty activating plurality in processing plural inanimate nouns in English. Therefore, there was expected to be a smaller singular-plural RT difference in processing inanimate nouns in the number judgment task. However, the results were opposite: Japanese learners of English demonstrated a larger interference effect for inanimate nouns.

This rather contradictory result may be due to a collective interpretation derived from the combination of animate nouns and the plural morpheme. In fact, among the 39 animate nouns included in the data analysis, only seven were non-human animate nouns (animals); all others were human nouns. As reviewed in Section 2.2, Japanese *-tachi* could induce a collective interpretation rather than plurality (Nakanishi & Tomioka, 2004). If the English plural morpheme *-s* is associated with the Japanese plural marker *-tachi*, the plural form of human animate nouns might have been more likely to be interpreted as a group. This would have created less conflict in the number judgment task. Nonetheless, a note of caution is due here since the interaction term in the final GLMM model was not significant, and the p-value of the simple main effect test was close to the level of .05.

Another, more likely explanation of the lack of an animacy effect in processing plural nouns is that Japanese does accept inanimate plural nouns, as pointed out by

Murahata (2019). As mentioned in the background section, it is not entirely impossible to attach plural markers to inanimate nouns in Japanese, though it is not very common. Thus, the learners succeeded in accessing the plurality inherent in the plural morpheme attached to inanimate nouns as well as animate nouns. However, given that the processing of inanimate plural nouns is not common in Japanese, the participants could have had difficulty in processing inanimate plural nouns, resulting in longer RTs.

Limitations and Future Directions

Although the present study found a limited influence of animacy on the acquisition of plural morphemes, several limitations should be addressed in future research. The most important lies in the fact that the present study did not include L2 learners whose first language is not Japanese. Since the Animacy Hierarchy predicts a variation in the pluralization of nouns between languages, future research needs to compare L2 learners from different L1 backgrounds. As reviewed earlier in this paper, for example, in Russian, both animate and inanimate nouns can be pluralized, as in English. Therefore, Russian learners are expected to show no difficulty acquiring the plural morpheme in English, as demonstrated by Jiang et al. (2011). If Chinese learners of English whose L1 has optional plural marker -*men* that cannot be attached to animate and inanimate nouns, show a similar tendency to Russian learners of English and the Japanese participants in the current study, we could make a stronger argument that the animacy of nouns does not play a significant role in the acquisition of the plural morpheme in English. Therefore, a further study could assess the possible influence of the mass/count distinction on the number judgment task used in this study.

Another limitation is the interpretation of the reversed animacy effect. As discussed in the previous section, it was unclear why the singular-plural difference was smaller for animate nouns than inanimate nouns, contrary to the prediction based on the Animacy Hierarchy. Possibly related to this issue is the proficiency of the learners. The participants of the current study have the mean level of B1 in English. For lower proficiency learners who have not fully acquired the number feature might have difficulty activating plurality in processing inanimate nouns, thereby demonstrating animacy effect. In contrast, higher proficiency learners might activate plurality from animate and inanimate nouns to the same extent. As a result, for both animate and inanimate nouns, there would be significant RT difference between singulars and plurals. Since the limited number of participants in various proficiency levels, the present study could not investigate the effect of proficiency in the activation of plurality in different types of nouns. This is an important issue for future research.

Lastly, further studies are necessary to closely investigate the possible (dis)association of collective meaning and plural meaning for Japanese learners of English. Lardiere's (2009, 2010) feature reassembly hypothesis might be a good starting point to further clarify the relationship among animacy, plurality, and collectivity.

Supplementary Material

All the raw data, test sentences, and data analytic procedures are available at Open Science Framework.

(https://osf.io/chnrj/?view_only=012ea92af42f41679bd6011f54ef00e4)

References

Author. (***). XXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Asahara, M., Maekawa, K., Imada, M., Kato, S., & Konishi, H. (2014). Archiving and analysing techniques of the ultra-large-scale web-based corpus project of NINJAL, Japan. *Alexandria*, *25*(1–2), 129–148. https://doi.org/10.7227/ALX.0024

Bailey, N., Madden, C., & Krashen, S. D. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning*, *24*(2), 235–243. https://doi.org/10.1111/j.1467-1770.1974.tb00505.x

Bannai, M. (2011). The nature of variable sensitivity to agreement violations in L2 English. *EUROSLA Yearbook*, *11*, 115–137. doi:10.1075/eurosla.11.08ban

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2021). *lme4: Linear mixed-effects models using Eigen and S4* [R package version 1.1-27.1]. Available at: https://cran.r-project.org/package=lme4

Berent, I., Pinker, S., Tzelgov, J., Bibi, U., & Goldfarb, L. (2005). Computation of semantic number from morphological information. *Journal of Memory and Language*, *53*(3), 342–358. https://doi.org/10.1016/j.jml.2005.05.002

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Choi, S. H., & Ionin, T. (2021). Plural marking in the second language: Atomicity, definiteness, and transfer. *Applied Psycholinguistics*, *42*(3), 549–578. https://doi.org/10.1017/S0142716420000569

Choi, S. H., Ionin, T., & Zhu, Y. (2018). L1 Korean and L1 Mandarin L2 English

learners' acquisition of the count/mass distinction in English. *Second Language*

Research, *34*(2), 147–177. https://doi.org/10.1177/0267658317717581

Corbett, G. G. (2000). *Number*. Cambridge University Press.

Corbett, G. G. (2012). *Features*. Cambridge University Press.

Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language

acquisition. *Language Learning*, *24*(1), 37–53.

https://doi.org/10.1111/j.1467-1770.1974.tb00234.x

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A

psychometric study. *Studies in Second Language Acquisition*, *27*(2), 141–172.

https://doi.org/10.1017/S0272263105050096

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The

differential contributions of implicit and explicit knowledge. *Applied Linguistics*,

27(3), 431–463. https://doi.org/10.1093/applin/aml022

Hakuta, K. (1976). A case study of a Japanese child learning English as a second

language. *Language Learning*, *26*(2), 321–351.

https://doi.org/10.1111/j.1467-1770.1976.tb00280.x

Haspelmath, M. (2013). Occurrence of nominal plurality. In M. S. Dryer & M.

Haspelmath (Eds.), *The world atlas of language structures online.* Leipzig: Max

Planck Institute for Evolutionary Anthropology. Available at

http://wals.info/chapter/34

Iljic, R. (1994). Quantification in Mandarin Chinese: Two markers of plurality.

Linguistics, *32*(1), 91–116. https://doi.org/10.1515/ling.1994.32.1.91

Inagaki, S. (2014). Syntax–semantics mappings as a source of difficulty in Japanese

speakers' acquisition of the mass–count distinction in English. *Bilingualism:*

Language and Cognition, *17*(3), 464–477.

https://doi.org/10.1017/S1366728913000540

Ji, J., & Liang, M. (2018). An animacy hierarchy within inanimate nouns: English corpus

evidence from a prototypical perspective. *Lingua*, *205*, 71–89.

https://doi.org/10.1016/j.lingua.2017.12.017

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied*

Psycholinguistics, *25*(4), 603–634. https://doi.org/10.1017/S0142716404001298

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language

learning. *Language Learning*, *57*(1), 1–33.

https://doi.org/10.1111/j.1467-9922.2007.00397.x

Jiang, N., Hu, G., Chrabaszcz, A., & Ye, L. (2017). The activation of grammaticalized

meaning in L2 processing: Toward an explanation of the morphological

congruency effect. *International Journal of Bilingualism*, *21*(1), 81–98.

https://doi.org/10.1177/1367006915603823

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological

congruency and the acquisition of L2 morphemes. *Language Learning*, *61*(3),

940–967. https://doi.org/10.1111/j.1467-9922.2010.00627.x

Kato, S (2006). Nihongo no fukusuukei: Futatsu no fukusuu to shugo ninchi [Japanese

plural: Two types of plural and cognition of group]. *Kokugo Kokubun Kenkyu,*

130, 62–48.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second

    language acquisition. *Second Language Research*, *25*(2), 173–227.

    https://doi.org/10.1177/0267658308100283

Lardiere, D. (2010). Feature assembly in second language acquisition. In J. Liceras, H.

    Zobl, & H. Goodluck (Eds.), *The role of formal features in second language*

    acquisition (pp. 107–140). Routledge.

Larsen-Freeman, D. E. (1975). The acquisition of grammatical morphemes by adult ESL

    students. *TESOL Quarterly*, *9*(4), 409–419. https://doi.org/10.2307/3585625

Lempert, H. (2016). Late L2ers can acquire grammatical features that do not occur in

    their L1: Evidence from the effect of animacy on verb agreement in L1 Chinese.

    Memory & Cognition, *44*(4), 538–553.

    https://doi.org/10.3758/s13421-015-0583-6

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means* [R

    package version 1.7.0]. Available at:

    https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Li, Y.-H. A. (1999). Plurality in a classifier language. *Journal of East Asian Linguistics*,

    8(1), 75–99.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear

    mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171.

    https://doi.org/10.3389/fpsyg.2015.01171

Luk, Z. P., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes

    impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles,

    and possessive 's. *Language Learning*, *59*(4), 721–754.

    https://doi.org/10.1111/j.1467-9922.2009.00524.x

Morita, Y. (1980). *Kiso Nihongo 2* [Basic Japanese 2]. Kadokawa-shoten.

Murahata, Y. (2019). Nihongo no fukusuu hyoshiki ni mirareru eigo no eikyou: *-tachi* ga hyousyou suru kobetsuka no kanten kara [Influence of English on Japanese plural marker: From the viewpoint of individuation represented by *-tachi*]. *Kyouiku Kagaku Ronsyu, 6*, 15–27.

Murakami, A., & Alexopoulou, T. (2016). L1 influence on the acquisition order of English grammatical morphemes. *Studies in Second Language Acquisition*, *38*(3), 365–401. https://doi.org/10.1017/S0272263115000352

Nakanishi, K., & Tomioka, S. (2004). Japanese plurals are exceptional. *Journal of East Asian Linguistics*, *13*(2), 113–140. https://doi.org/10.1023/B:JEAL.0000019058.46668.c1

Patson, N. D., & Warren, T. (2010). Evidence for distributivity effects in comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 782–789. https://doi.org/10.1037/a0018783

Shibuya, M., & Wakabayashi, S. (2008). Why are L2 learners not always sensitive to subject-verb agreement? *EUROSLA Yearbook*, *8*, 235–258. doi:10.1075/eurosla.8.13shi

Song, Y. (2015). L2 Processing of plural inflection in English: L2 processing of plural inflection in English. *Language Learning*, *65*(2), 233–267. https://doi.org/10.1111/lang.12100

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, *38*(5), 1229–1261. https://doi.org/10.1017/S014271641700011X

Trenkic, D., Mirkovic, J., & Altmann, G. T. M. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition*, *17*(2), 237–257. https://doi.org/10.1017/S1366728913000321

Tsang, A. (2017). Judgement of countability and plural marking in English by native and non-native English speakers. *Language Awareness*, *26*(4), 343–359. https://doi.org/10.1080/09658416.2017.1410553

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests. *Studies in Second Language Acquisition*, *39*(01), 59–95. https://doi.org/10.1017/S0272263115000455

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Wen, Z., Miyao, M., Takeda, A., Chu, W., & Schwartz, B. D. (2010). Proficiency effects and distance effects in nonnative processing of English number agreement. In K. Franich, K. M. Iserman, & L. L. Keil (Eds.), *Proceedings of the 34th Annual Boston University Conference on Language Development* (pp. 445–456). Cascadilla Press.