Agreement affects the interpretation of null arguments in semi-artificial Japanese

Naho Orita (orita@waseda.jp)
Faculty of Science and Engineering, Waseda University, Japan

Daiko Takahashi (daiko@tohoku.ac.jp)
Graduate School of International Cult ural Studies, Tohoku University, Japan

Hiromu Sakai (hsakai@waseda.jp)
Faculty of Science and Engineering, Waseda University, Japan

Abstract

The availability of sloppy interpretation for null arguments differs across languages. This difference is a challenge for child learners because they are unlikely to receive sufficient input that provides clear evidence about the available interpretation. Previous theoretical work suggests that knowing the presence/absence of agreement could help this learning problem. While languages like Japanese that lack agreement allow argument ellipsis (hence the sloppy reading is available), languages with a rich agreement system like Spanish do not. This study explores the utility of this correlation as a cue for learners to infer the available interpretation of a null argument. We show that Japanese adults who learned semi-artificial Japanese that has object-verb agreement are more likely to accept the strict reading than the ones who learned only an artificial singular/plural marker attached to an object. We also find that the way of presenting visual stimuli largely affects participants' interpretation of null arguments. Our results provide evidence that agreement may play a role in learning the interpretation of null arguments.

Keywords: null arguments, argument ellipsis, artificial language learning, Japanese

Introduction

Given the preceding sentence (1a), a null object construction in Japanese (1b) has two interpretations: *Ken respects Taro's mother* and *Ken respects his own mother*, the so-called strict reading and sloppy reading respectively.

(1) a. Taro-wa zibun-no-hahaoya-wo sonkeishite-iru.
Taro-TOP self-POSS-mother-ACC respect-PRES
'Taro respects self's mother.'

b. Ken-mo sonkeishite-iru.
Ken-also respect-PRES
'Ken respects [e] too.'

On the other hand, a null subject in Spanish (2b) has only the strict reading given the preceding sentence (2a): *Juan believes that María's proposal will be accepted* (the data cited from Oku, 1998).

(2) a. Maria cree que su propuesta sera aceptada.
Maria believes that her proposal will-be accepted
'Maria believes that her proposal will be accepted.'

b. Juan tambien cree que sera aceptada.
Juan also believes that will-be accepted
'Juan also believes that [e] will be accepted.'

The availability of sloppy reading has been attributed to the theoretical status of those null elements. On one hand, the argument ellipsis analysis proposes that a null element in languages like Japanese is a result of eliding the full-fledged DP in the corresponding grammatical position (Oku, 1998, inter alia). On the other hand, a null element in languages like Spanish has been analyzed as an empty pronoun which is base-generated in the corresponding position (Jaeggli & Safir, 1989, inter alia).

This cross-linguistic difference has been linked to the presence/absence of agreement between the argument and its associated functional head. Saito (2007) and Takahashi (2014, 2020) propose the *anti-agreement analysis* where languages that lack agreement (e.g., Japanese) allow argument ellipsis while languages with rich agreement (e.g., Spanish) do not. The latter null elements, which are accompanied by agreement, are considered as empty pronouns and hence they only yield the strict reading. This hypothesis has been confirmed in various languages. On one hand, a group of languages without agreement, such as Japanese, Korean, and Mongolian, allow both null subjects and null objects, and these null arguments allow the sloppy reading as well as the strict reading (Takahashi, 2007; Sakamoto, 2012). On the other hand, languages with a rich agreement system, such as Kaqchikel and Spanish, do not seem to allow the sloppy reading at all (Otaki, Sugisaki, Yusa, & Koizumi, 2013).[1] Interestingly, some languages that exhibit only (implicit) subject agreement, such as Turkish, Chinese, and Malayalam, allow the sloppy reading with null objects but not with the null subjects (Takahashi, 2014, 2020; Sato & Karimi, 2016).

The cross-linguistic correlation between agreement and argument ellipsis would become crucial for children acquiring language because the null elements are unpronounced, and thus they themselves do not show any clear distributional evidence that would lead learners to identify the available interpretations in their language. Crucially, the availability of sloppy reading is arguably determined by grammar but not solely by context. In the case of Japanese, the occurrences of anaphor *zibun* (self), which is a component of the elided NP as in (1a), is quite rare in child-directed speech, and the most frequent use of *zibun* in the input is as a second-person indexical (Orita, Ono, Feldman, & Lidz, 2021). Nevertheless,

[1] Some exceptions in Spanish are reported in Duguine (2014).

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

Table 1: Prediction summary

Language	Agreement	Strict	Sloppy
Semi-Japanese	Absent	True	True
	Present	True	False ↑

Japanese preschool children demonstrate adult-like interpretations of null arguments (Sugisaki, 2007, 2018).

Given this gap between input and what children know and the fact that agreement errors rarely occur in children learning languages with a rich agreement system (Hyams, 2002), we argue that the presence/absence of agreement plays a role in identifying possible interpretations of null arguments in a language. At this stage, we refrain from arguing whether and what kind of innate linguistic knowledge might contribute to this learning problem, but as a first step toward exploring such a question, we experimentally investigate the influence of agreement on the interpretations of null arguments.

Our experiments adapt an artificial language learning method called the extrapolation paradigm (Culbertson & Adger, 2014). A growing body of research using this paradigm has demonstrated a relation between typology and natural language acquisition (e.g., Culbertson, Smolensky, & Legendre, 2012; Martin, Ratitamkul, Abels, Adger, & Culbertson, 2019; Maldonado & Culbertson, 2022). In this paradigm, learners are first exposed to input that does not contain critical evidence about the target knowledge and then asked to generalize from that ambiguous evidence. This method allows us to test whether and how learners' inferences are guided by a particular knowledge or bias and to control the influence of a particular factor of interest. We adapt this method to see whether learners' interpretations of null arguments, namely strict and sloppy reading, are influenced by the presence/absence of agreement.

Given the anti-agreement hypothesis, we predict that Japanese speakers who learned a semi-artificial Japanese with agreement are more likely to reject the sloppy reading (Table 1 for summary). This paper reports two experiments that test this prediction. We show that Japanese adult speakers who learned artificial object-verb agreement are more likely to accept the strict reading than the ones who learned only artificial singular/plural markers attached to the object. Although the acceptance rates of the sloppy reading did not change regardless of the presence of agreement, which is against our prediction, the results still demonstrate the influence of learned agreement on the interpretation of null arguments. Our secondary finding is that the strict reading is susceptible to the way of presenting visual stimuli, suggesting that special care needs to be taken to ensure the intended reading.

One previous work that should be noted here is an experimental study with Spanish and German learners of Japanese. Yamada and Miyamoto (2002) tested the interpretations of Japanese null subjects and objects by native Spanish (pro-drop) and German (non-pro-drop) speakers who learn Japanese as a foreign language. They found that Spanish learners of Japanese were more likely to allow the sloppy interpretation than German learners of Japanese. Though their research question and method are not identical to our study, we think their experiments are relevant to our interest in that the anti-agreement hypothesis would also predict that Spanish speakers who learned a semi-artificial Spanish without agreement are more likely to accept the sloppy reading. We believe our experiments using an artificial language with monolingual speakers provide a more controlled experimental setting and hence more reliable results than collecting judgments from second language learners who are known to have a large variation in proficiency (Skehan, 1991).

Experiment 1

Methods

We adapt an artificial language learning method called the extrapolation paradigm (Culbertson & Adger, 2014) to investigate the influence of agreement on learning the interpretation of null arguments. Learners in this experiment are exposed to semi-Japanese that only differs from Japanese with respect to agreement. While Japanese lacks agreement, the semi-Japanese has object-verb agreement.

We choose artificial object-verb agreement to minimize complexity in the semi-artificial language. If subject-verb agreement is used instead, then the sentence to test the sloppy/strict reading would become more complex since it requires an embedded clause as in (2). This increases the complexity both in the input language and the construction of contexts to elicit the target reading. Thus, we choose the null object construction.

In the training, learners are exposed to simple SOV sentences from semi-Japanese. At test, learners infer the interpretation of null arguments in the semi-Japanese. The critical point here is that learners are trained on input that includes no evidence about the interpretation of null arguments, namely sloppy reading and strict reading. Their judgments indicate whether the learned agreement influences their interpretations of null arguments. If there is a relation between agreement and the interpretation of null arguments, learners who are exposed to the input that contains artificial agreement (Agreement condition) would be less likely to accept the sloppy reading items. We compare these learners with a control group where learners are only exposed to artificial singular/plural markers (i.e., no agreement) and examine the differences between them.

Participants Participants were 42 undergraduate and graduate students at Waseda University, Japan. All were native Japanese speakers and at least 18 years old. They were recruited via the Waseda University's student job center. Participants were paid ¥1,000 for an hour long experimental session, including instruction, a short break, and debriefing.

Input language The lexicon of this experiment consists of: 4 verbs (*keru* 'kick', *osu* 'push', *motsu* 'have', and *tataku*

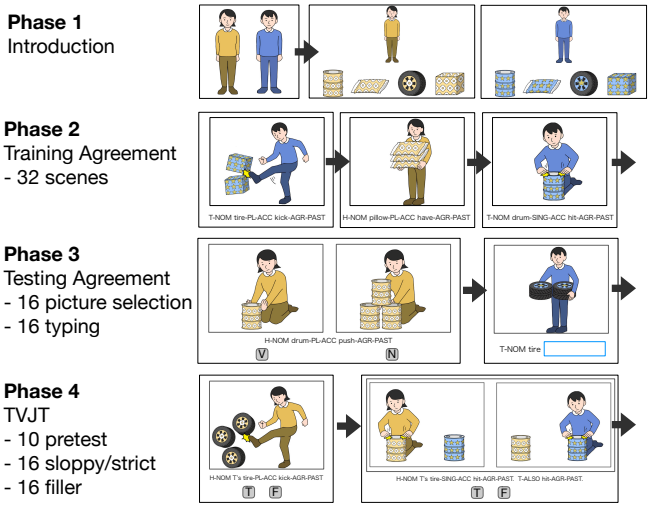

Figure 1: Example visual stimuli in Experiment 1

'hit'), 2 subject nouns (proper names *Taro* and *Hanako*), 4 object nouns (*hako* 'box', *makura* 'pillow', *taiya* 'tire', *kan* 'drum'), 2 nonce post-nominal singular/plural markers (*-pi* for singular and *-pepu* for plural), and 2 nonce post-verbal agreement markers (*-pa* for singular and *-po* for plural). Word order is always SOV (agent-patient-verb). All the stimuli are displayed in Katakana script with audio. The auditory stimuli were recorded by a female native Japanese speaker. The following shows sample sentences used in the training phase of the Agreement condition (3) and the Control condition (4). The crucial difference is that the lexicon of the Control condition does not include agreement markers but only the singular/plural markers attached to the object.

(3) Agreement condition

a. Hanako-ga hako-**pi**-wo tataki-**pa**-ta.
Hanako-NOM box-**SING**-ACC hit-**AGR**-PAST
"Hanako hit a box."

b. Taro-ga taiya-**pepu**-wo keri-**po**-ta.
Taro-NOM tire-**PL**-ACC kick-**AGR**-PAST
"Taro kicked tires."

(4) Control condition

a. Hanako-ga hako-**pi**-wo tatai-ta.
Hanako-NOM box-**SING**-ACC hit-PAST
"Hanako hit a box."

b. Taro-ga taiya-**pepu**-wo ket-ta.
Taro-NOM tire-**PL**-ACC kick-PAST
"Taro kicked tires."

Experimental procedure and materials Participants were tested individually in a quiet room. All instructions were provided in Japanese. The session consists of the following 4 phases. Figure 1 summarizes sample visual stimuli. The experiment was conducted using PsychoPy (Peirce, 2007).[2]

Phase 1 Introduction Participants are first introduced to 2 subject nouns and 4 object nouns in isolation. They are told that Hanako's objects are yellow and flower printed and that Taro's objects are blue and star printed. After the introduction, they do a noun-selection task (6 trials).

Phase 2 Training Agreement At the beginning of the training trials, participants are told that the language is similar to Japanese but they would notice some differences. During training, participants in the Agreement group are exposed to input that exhibits object-verb agreement and participants in the Control group are exposed to input that only contains the artificial singular/plural markers (32 trials each). On each training trial, participants see an image along with its corresponding script (as in (3) or (4)) and audio.

Phase 3 Testing Agreement Participants are first tested on their understanding of object-verb agreement (Agreement condition) or singular/plural markers (Control condition) in an image-selection task. On each trial, participants see an array of two images with a description in text accompanying audio. They must select the image that correctly represents the description. Feedback sound is provided after selecting the image. Each foil image represents the same event with wrong object(s), e.g., singular instead of plural. Total 16 trials (8 singular correct and 8 plural correct items) are randomly presented.

After the image selection task, participants take a typing task to be tested on their production of object-verb agreement or singular/plural markers. On each trial, participants see an image with an incomplete sentence (e.g., 'Taro-NOM box ') and are asked to type in the remaining words. Total 16 trials (8 singular and 8 plural items) are randomly presented.

Phase 4 TVJT To provide an independent measure of participants' understanding of the task, a pretest session (a simple True/False judgment task) is conducted before the last trials. The session consists of 10 trials (5 true and 5 false items), e.g., subject/object/possessor mismatch, verb mismatch, and agreement/number mismatch.

Participants then perform the Truth-Value Judgment task (Crain & Thornton, 1998). On each trial, a set of two images with sentences describing a short context is presented (see Figure 1 Phase 4), followed by a target sentence that contains a null object construction. Participants are asked to judge whether the target sentence truthfully described the given context by clicking True or False. These sentences are created by using the lexicon presented in the preceding phases. In addition to this lexicon, two Japanese words, *zibun* 'self' and *-mo* 'also,' are also used. The following (5) and (6) show sample test and filler items.

The 64 possible test items (4 verbs × 4 object nouns × 2 agreement markers × 2 context types = 64 items) are created and then assigned to four lists using a Latin square design,

[2]Sample experimental materials and the data analysis plan for Experiment 1 and 2 are available at https://osf.io/2f9ya/.

so that each list contains the same frequency of verbs, nouns, markers, and context types. Each list results in 8 sloppy-true and 8 strict-true items. Likewise, filler items (all False; two context types as in (6)) are assigned to four lists. Total 32 items are randomly presented to each participant.

(5) Sample test item

Hanako-ga zibun-no-kan-pi-wo
Hanako-NOM self-POSS-drum-SING-ACC
tataki-pa-ta. Taro-mo tataki-pa-ta.
hit-AGR-PAST Taro-also hit-AGR-PAST
"Hanako hit self's drum. Taro hit [e] too."

- Sloppy-true: Hanako hit Hanako's drum. Taro also hit Taro's drum.
- Strict-true: Hanako hit Hanako's drum. Taro also hit Hanako's drum.

(6) Sample filler item

Taro-ga zibun-no-makura-pepu-wo oshi-po-ta.
Taro-NOM self-POSS-pillow-PL-ACC push-AGR-PAST
Hanako-mo oshi-po-ta.
Hanako-also push-AGR-PAST
"Taro pushed self's pillows. Hanako pushed [e] too."

- Null-object mismatch: Taro pushed Taro's pillows. Hanako also pushed Hanako's boxes.
- Overt-object mismatch: Taro pushed Taro's tires. Hanako also pushed Hanako's boxes.

Results

Participants who made more than 40% errors in the pretest session were excluded from the analysis (N=2), resulting in 21 participants in the Agreement condition and 19 participants in the Control condition.

We analyzed the number of "False" responses to the filler items and "True" responses to the test items. Figure 2 and Figure 3 summarize the mean percentages of these items by condition. The participants' individual means are shown as transparent circles and triangles. The overall means are shown as opaque with standard errors.

We constructed two separate logistic mixed-effects models, one for predicting the rejection of filler items and one for predicting the acceptance of the test items. The models include Condition (Agreement and Control) and Context (Null-object mismatch and Overt-object mismatch in the filler items and Sloppy-true and Strict-true in the test items) as fixed effects. The models also include their interaction and random intercepts for participants and items.[3]

Overall, participants correctly rejected almost all filler items as false (Figure 2), except the Null-object mismatch

[3]The models with maximal random effects structure failed to converge, so we report the models without the random slopes.

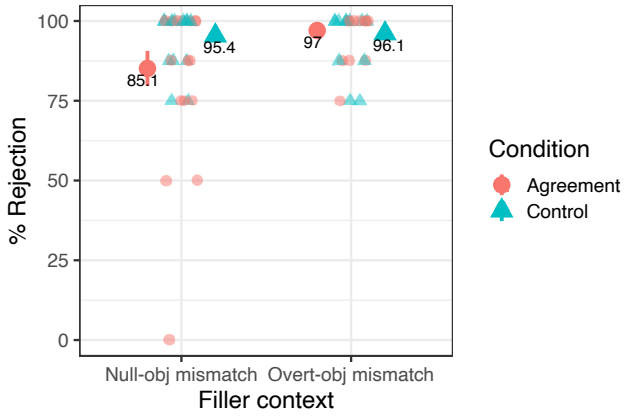

Figure 2: Experiment 1 rejection rates for filler items

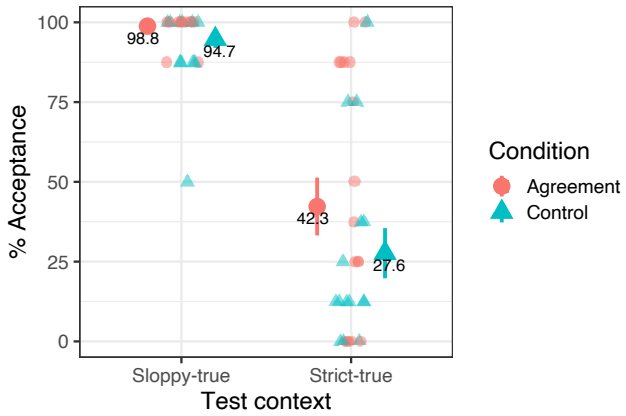

Figure 3: Experiment 1 acceptance rates for test items

items in the Agreement condition, yielding a significant interaction between Condition and Context ($\beta = -2.16$, $SE = 0.86$, $z = -2.50$ $p = 0.01$). This type of filler presents a mismatch between the object in the image and the referent of the null object in the sentence, so it should be more difficult than the other filler that presents a mismatch with the overt object in the sentence. Besides, participants who learned the artificial agreement would have borne more processing cost. The combination of the task difficulty and processing cost might have affected the performance of the Null-object mismatch in the Agreement condition.

Table 2: Experiment 1: Model output summary (test items)

	β	SE	z	$Pr(>\|z\|)$
(Intercept)	4.12	0.72	5.72	<0.001
Condition (Agreement)	2.41	1.22	1.98	0.048
Context (Strict-true)	-5.55	0.61	-9.17	<0.001
Condition × Context	-1.59	1.12	-1.42	0.16

As shown in Figure 3 and Table 2, we found significant effects of Condition and Context but no significant interaction between them. These results indicate that participants in the Agreement condition are more likely to accept the test items in general and that both the Agreement and Control groups are less likely to accept the strict reading. Contrary to our prediction, the acceptance rate of the sloppy reading did not decrease in the Agreement group. Both groups similarly accepted the sloppy reading. We will discuss this point later in the Discussion. Table 2 (as well as Table 3 of Experiment 2) also indicates that the random intercept is significant. The analysis revealed that there is a variation in participants' responses. We return to this issue in the Discussion.

Although there was no significant interaction between Condition and Context, we observe a tendency that participants who learned the agreement are more likely to accept the strict reading. Moreover, the overall low acceptance rates in the strict reading is indicative in that some unforeseen factor might have affected this result because the strict reading in Japanese is normally allowed as shown in (1). We suppose that displaying the set of two separate images side by side at the same time may make it harder for participants to arrive at the intended reading because the situation could be interpreted such that each character is carrying out an action toward each object in parallel, but not toward a single identical object. For example, in the Strict-true context (5) with the two separate images, the test sentence containing a null argument could be interpreted as *Taro also hit Hanako's box which is different from the box that Hanako hit*. This kind of ambiguity could have suppressed the strict reading. The following Experiment 2 addresses this potential issue by changing the visual stimuli from still image to movie.

Experiment 2

To solve the potential problem identified in Experiment 1, all the still images used in Phase 2, 3, and 4 of Experiment 1 are replaced with movies in Experiment 2. For example, in the case of Strict-true context in (5), Experiment 1 displayed two separate still images at the same time: the left side image describes that Hanako hit Hanako's drum and the right side image describes that Taro hit Hanako's drum. Instead, Experiment 2 shows a sequence of events in one movie: Hanako appeared in the scene, hit Hanako's drum, and left the scene; then Taro appeared in the same scene, hit Hanako's drum, and left the scene.

Since playing all the movies takes much longer time than showing the still images, the number of trials in each phase is reduced as follows: Phase 2 (training agreement) 16 trials (8 singular and 8 plural); Phase 3 (testing agreement) 16 image selection trials and 8 typing trials; Phase 4 (TVJT) 10 pretest trials, 8 test items (4 sloppy and 4 strict), and 8 filler items. The basic structure of the experiment and input language are same as Experiment 1.

Experimental procedure Due to the COVID-19 restrictions, we opted to conduct Experiment 2 online using the Gorilla Experiment Builder[4] (Anwyl-Irvine, Massonie, Flitton, Kirkham, & Evershed, 2019). We used this platform to create and host the experiment. During the experiment, each participant was individually monitored by an experimenter using a video conferencing platform. In Phase 2 to 4, participants were allowed to play each movie as many times as they want, but almost none of them did so.

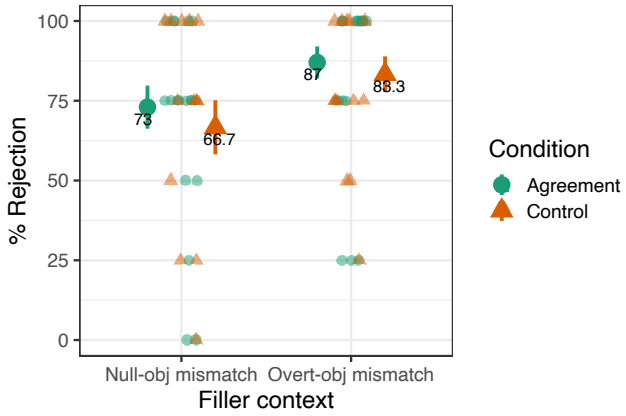

Figure 4: Experiment 2 rejection rates for filler items

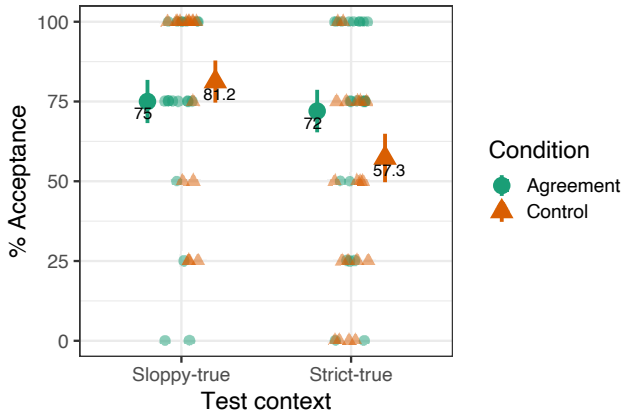

Figure 5: Experiment 2 acceptance rates for test items

Participants Participants were 66 undergraduate and graduate students at Waseda University, Japan. All were native Japanese speakers and at least 18 years old. The recruitment procedure is same as Experiment 1. Participants who had trouble with their online experimental setting (e.g., unstable WiFi connection and audio issues) were excluded, resulting in 52 participants for analysis.

Results

We conducted the same analysis as Experiment 1. Participants who made more than 40% errors in a pretest session were excluded from the analysis (N=3). This resulted in 25 participants in the Agreement condition and 24 participants

[4] www.gorilla.sc

Table 3: Experiment 2: Model output summary (test items)

	β	*SE*	*z*	*Pr*(> |*z*|)
(Intercept)	1.60	0.31	5.18	<0.001
Condition (Agreement)	-0.40	0.41	-0.99	0.32
Context (Strict-true)	-1.27	0.35	-3.64	<0.001
Condition × Context	1.11	0.48	2.31	0.021

in the Control condition.

As shown in Figure 4, the overall rejection rates for filler items were lower than Experiment 1. We think this is because the movie stimuli require greater working memory than still images. Although participants were allowed to repeat the movie in each trial, almost all of them watched it just once. In contrast to Experiment 1, there was a significant fixed effect of Context ($\beta = -1.72$, $SE = 0.52$, $z = -3.33$, $p < 0.001$) but no interaction between Condition and Context in the filler items. These results suggest that participants were more likely to (wrongly) accept the Null-object mismatch items than the Overt-object mismatch items, regardless of the condition. This is reasonable because the Null-object mismatch items are essentially more difficult.

Figure 5 shows that the acceptance rates of the strict reading across two conditions are higher than the ones in Experiment 1. This shows that changing the visual stimuli affects the interpretation of strict reading and thus the identicality of the referent is crucial to make the strict reading available.

As shown in Table 3, we found a significant fixed effect of Context and a significant interaction between Condition and Context. There is a difference between Sloppy-true and Strict-true in the Control condition, while the Agreement condition does not show such difference. In other words, participants are equally likely to accept the sloppy and strict readings in the Agreement condition while participants in the Control condition are not. These results suggest that there is a significant influence of learned agreement on the interpretations of null objects. Though these results do not support our first prediction that Japanese speakers who learned agreement are more likely to reject the sloppy reading, the difference in the strict reading demonstrates the influence of the presence of agreement in the interpretation of null arguments.

Discussion

This study examined the influence of agreement on the interpretation of null arguments. We found that Japanese adult speakers who learned artificial object-verb agreement are more likely to accept the strict reading than the ones who only learned artificial singular/plural markers attached to the object. Our results provide evidence that agreement may play a crucial role in learning the interpretation of null argument.

The difference in the acceptance of strict reading in Experiment 2 may indicate that the presence of agreement facilitates the learners to interpret the null argument as a zero pronoun. Previous theoretical work suggests that there could be two types of null elements in a language (Oku, 1998; Takahashi, 2013). The zero pronoun option might have become salient after learning the agreement.

Contrary to our prediction, the acceptance rates of the sloppy reading in the Agreement condition did not change in both Experiment 1 and 2. We argue that this is due to an influence of L1 Japanese. It would be difficult for Japanese adult speakers to cancel the sloppy reading which they have already acquired, particularly in the semi-Japanese of this experiment, which was identical to Japanese except the object-verb agreement. Crucially, we used the anaphor *zibun* as it is, which can be interpreted as both sloppy and strict reading. To reduce the influence of L1 Japanese, we could create a more artificial language that is substantially different from Japanese, along with a long-term and complex training phases. However, it is not entirely clear how the grammatical knowledge of this kind of anaphor could be trained without using the null construction, which is the phenomenon of interest in this study.

The acceptance rates of the strict reading in Experiment 2 largely increased from Experiment 1. This difference indicates that the strict reading is sensitive to the identicality of the referent of the null argument. The sloppy/strict interpretation tends to be elicited based on linguistic context (text), particularly when data is collected from adult speakers (e.g., Han, Kim, Moulton, & Lidz, 2020). Our results suggest that such judgments would become more robust if this kind of issue is clearly controlled in the context.

In both Experiment 1 and 2, there was variability in participants' responses. We will collect more data to confirm the reliability of observed differences. Another potential problem is that there is a possibility that participants learned or interpreted the artificial agreement as something else, such as a clitic pronoun. It would be important to conduct an independent test to rule out this possibility. Alternatively, we could confirm the results of this study by testing our hypothesis with Spanish speakers. The prediction is opposite to that of in Japanese. If agreement plays a role in identifying possible interpretations of null arguments in a language, Spanish speakers who learned a semi-artificial Spanish without agreement would be more likely to accept the sloppy reading.

Acknowledgments

We thank anonymous reviewers for their helpful suggestions and comments. We also thank Sawano Endo for creating the illustrations and movies used in this study. This research was supported by JSPS KAKENHI 20K00615.

References

Anwyl-Irvine, A., Massonie, J., Flitton, A., Kirkham, N., & Evershed, J. (2019). Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods*. doi: 10.3758/s13428-019-01237-x

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. MIT Press.

Culbertson, J., & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences*, *111*(16), 5842–5847. doi: 10.1073/pnas.1320525111

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*, 306–329.

Duguine, M. (2014). Argument ellipsis: a unitary approach to pro-drop. *The Linguistic Review*, *31*(3-4), 515–549.

Han, C.-h., Kim, K.-m., Moulton, K., & Lidz, J. (2020). Null objects in Korean: Experimental evidence for the argument ellipsis analysis. *Linguistic Inquiry*, *51*(2), 319–340.

Hyams, N. (2002). Clausal Structure in Child Greek: A reply to Varlokosta, Vainikka and Rohrbacher and a reanalysis. *The Linguistic Review*, *19*(3), 225–269. doi: doi:10.1515/tlir.2002.001

Jaeggli, O., & Safir, K. J. (1989). The Null Subject Parameter and Parametric Theory. In O. A. Jaeggli & K. J. Safir (Eds.), *The null subject parameter* (pp. 1–44). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-009-2540-3_1

Maldonado, M., & Culbertson, J. (2022). Person of Interest: Experimental Investigations into the Learnability of Person Systems. *Linguistic Inquiry*, *53*(2), 295-336. doi: $10.1162/\text{ling}_{a0}0406$

Martin, A., Ratitamkul, T., Abels, K., Adger, D., & Culbertson, J. (2019). Cross-linguistic evidence for cognitive universals in the noun phrase. *Linguistics Vanguard*, *5*(1). doi: doi:10.1515/lingvan-2018-0072

Oku, S. (1998). *A theory of selection and reconstruction in the minimalist perspective*. Unpublished doctoral dissertation, University of Connecticut.

Orita, N., Ono, H., Feldman, N. H., & Lidz, J. (2021). Japanese children's knowledge of the locality of zibun and kare. *Language Acquisition*, *28*(4), 327-343. doi: 10.1080/10489223.2021.1899181

Otaki, K., Sugisaki, K., Yusa, N., & Koizumi, M. (2013). The parameter of argument ellipsis: The view from Kaqchikel. *Studies in Kaqchikel grammar*, *8*, 153–162.

Peirce, J. W. (2007). PsychoPy –Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Saito, M. (2007). Notes on East Asian argument ellipsis. *Language research*, *43*(2), 203–227.

Sakamoto, Y. (2012). *A study of sluicing and cleft in Mongolian: A comparison with Japanese*. Unpublished master's thesis, Tohoku University.

Sato, Y., & Karimi, S. (2016). Subject-object asymmetries in Persian argument ellipsis and the anti-agreement theory. *Glossa: a journal of general linguistics*, *1*.

Skehan, P. (1991). Individual differences in second language learning. *Studies in second language acquisition*, *13*(2), 275–298.

Sugisaki, K. (2007). The configurationality parameter in the minimalist program: A view from child Japanese. In *Proceedings of the 31st annual boston university conference on language development* (pp. 597–608).

Sugisaki, K. (2018). Argument ellipsis in child Japanese revisited. In *Proceedings of the 10th workshop on altaic formal linguistics* (pp. 37–54).

Takahashi, D. (2007). Argument ellipsis from a crosslinguistic perspective: An interim report. In *GLOW in Asia VI.* The Chinese University of Hong Kong.

Takahashi, D. (2013). A note on parallelism for elliptic arguments. In *Formal Approaches to Japanese Linguistics* (Vol. 6, pp. 203–213).

Takahashi, D. (2014). Argument ellipsis, anti-agreement, and scrambling. In M. Saito (Ed.), *Japanese syntax in comparative perspective* (pp. 88–116). Oxford University Press Oxford.

Takahashi, D. (2020). Derivational argument ellipsis. *The Linguistic Review*, *37*(1), 47–74. doi: doi:10.1515/tlr-2019-2034

Yamada, K., & Miyamoto, Y. (2002). On the interpretation of null arguments in L2 Japanese by European non-pro-drop and pro-drop language speakers. *Journal of the European Second Language Association*, *1*(1), 73–89. doi: doi:10.22599/jesla.18