Influence of word-based prediction errors on syntactic priming

Manabu Arai Franklin Chang[1)]

Keywords: Syntactic priming; implicit learning; prediction error; passive structure

Abstract

Syntactic priming is a phenomenon where people tend to repeat syntactic structures that they have recently experienced. It has been argued to reflect error-based implicit learning of syntactic structures, where the error signal comes from word-level predictions in online comprehension. Previous studies, however, used sentence-level prediction errors measured offline and it is not yet clear whether syntactic priming indeed reflects word-level prediction errors. By making use of the ambiguous case marker *ni* in Japanese, that is used to denote either a goal role in the dative structure or an agent role in the passive structure, we manipulated structural expectations at the preverbal argument in prime sentences. Our results demonstrated a larger effect of syntactic priming when participants previously experienced the dative structure in filler items than when they did not experience any dative fillers. The finding is taken as the evidence for an influence of word-level prediction errors on syntactic priming.

1. Introduction

People tend to repeat syntactic structures that they have recently experienced.

1) Department of English Studies, Kobe City University of Foreign Studies

For example, one would be more likely to produce a passive sentence such as *The celebrity was embarrassed by the news anchor* after hearing a passive sentence like *The city was attacked by terrorists*. This phenomenon is called syntactic priming and occurs across utterances with different words and meanings (Bock, 1986). It has also been observed both in production and comprehension (Arai, van Gompel, & Scheepers, 2007; Branigan, Pickering, & McLean, 2005; Thothathiri & Snedeker, 2008). This suggests that the effect taps into a shared representation of syntactic structures that can adapt to recent exposure. Furthermore, it has been shown that priming can be quite long lasting (Bock & Griffin, 2000; Branigan, Pickering, Stewart, & McLean, 2000; Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008) and it is even possible for adults to learn novel sentence structures (Kaschak & Glenberg, 2004).

The persistence of priming has been used to argue that syntactic priming is a form of implicit learning rather than just transient activation of a particular syntactic structure (e.g., Bock & Griffin, 2000; Chang, Dell, Bock, & Griffin, 2000). To demonstrate this link, Chang, Dell, and Bock (2006) developed a model of syntactic priming using an error-based learning mechanism. The model would make predictions about the next word in sentences and the difference between the model’s prediction and the actual next word was called the error signal. The model was able to use this error signal to change the weights in the model so that the model was better able to predict the sentences that it experienced. Part of these changes involved changes to syntactic representations and this created syntactic priming if learning was left ON during the processing of the prime sentence. Support for error-based learning comes from studies that find that magnitude of priming is inversely correlated with structural preference of individual verbs (Barnolet & Hartsuiker, 2010, for production; Fine & Jaeger, 2013, for comprehension). When the verb’s preference mismatches the sentence structure, that mismatch should create larger prediction errors and yield more weight changes.

The idea of using the mismatch between expectations and actual inputs is a general idea that is also present in theories of language comprehension. One such approach is the surprisal theory (Hale, 2001; Levy, 2008). It assumes that processing cost is inversely correlated with the surprisal value, which is defined as the negative logarithm of the probability of a structure (*x*) given the preceding context as shown in (1).

(1) $$surprisal(x) = \frac{1}{\log probability(x)}$$

The surprisal value for a particular input would be large if it is unlikely on the basis of preceding input. In the classic example *The horse raced past the barn fell*, the surprisal for the word *fell* is very large because all the input preceding the verb supports the dominant (most frequent) main clause analysis (Bever, 1970). This explains the very large processing difficulty on encountering the verb *fell*, which is commonly referred to as a *garden-path* effect.

Surprisal has also been used to explain structural priming in a way that is similar to error-based learning (Fine & Jaeger, 2013; Fine, Jaeger, Farmer, & Qian, 2013; Hartsuiker et al., 2008). The main difference is that surprisal is estimated at a particular level (e.g., words, structure) and the effects of surprisal are also localized to that level. Surprisal accounts of sentence comprehension often use word-based surprisal, and surprisal accounts of syntactic priming use estimated surprisal of sentence structures in a particular corpus or norming study (Fine et al., 2013; Jaeger & Snider, 2013). For example, Fine et al. (2013) conducted a study using a self-paced reading task to investigate the influence of prediction errors on the subsequent processing of similar sentences using the sentences such as (2).

(2) The experienced soldiers warned about the dangers conducted the midnight

raid.

They found that the reading times for the ambiguous relative clause sentences like (2) decreased over the course of the experiment relative to those for the unambiguous counterpart (*The experienced soldiers who were told about the dangers conducted the midnight raid*) as predicted by the surprisal that they calculated. In their study, the surprisal value was estimated based on corpus statistics for individual verbs (i.e., how often a particular verb occured with the relative clause or the main clause structure occurs as a whole) as well as also local statistics (i.e., how many relative clause sentences had appeared at a given trial during the experiment). Error-based learning accounts of priming differ in that they assume that word-based error/surprisal is used to generate error/surprisal at the structural level. On this account, syntactic priming should be sensitive to word-based surprisal and that is not predicted in a pure surprisal account of priming.

Another issue with past studies is that the prediction errors based on the structural bias of a specific verb are likely to be lexically associated. For example, Jaeger and Snider (2013), who showed that their estimated surprisal can account for the priming effect in comprehension, calculated the lexical surprisal. That is, for a verb *give*, for example, the surprisal is estimated based on how often this particular verb appeared in the double object construction and in the prepositional object construction in the Switchboard corpus (See also Bernolet & Hartsuiker, 2010). Therefore, learning has to be at least partly lexically specific although the effect of the prediction error with a specific verb is generalized to a different verb (See Kaschak & Borreggine, 2008 for the finding that distributional information of syntactic structures for a particular verb caused little influence on a long-term structural priming effect). It is known that a priming effect tends to be stronger when the verb is repeated between prime and target than when it is not and previous research suggests that this lexical boost effect is likely to be attributed to explicit memory rather than error-based learning. In support of this, Hartsuiker et al. (2008) showed that a lexical boost effect is short-lived whereas an abstract priming effect is long-lasting and thus reflects error-based learning.

In this work, we hope to tease apart the word-level surprisal/prediction error effects in priming. In particular, we would like to see if these effects occur independently of verbs, which have many links to structural representations. Japanese provides a useful language to test these issues, because Japanese is a verb-final language, where structural processing must occur before the verb is encountered (Kamide, Altmann, & Haywood, 2003; Miyamoto, 2002). Syntactic priming has been found in both production as well as in comprehension with active/passive structures independently of the verb (e. g., Ishikawa, Arai, & Hirose, 2014 for production; Arai & Mazuka, 2014 for comprehension). To test for word-based prediction error, we manipulated the ambiguous case marker *ni* in the passive structure. In Japanese, the case marker *ni* is used not only to mark an agent role in the passive structure as in (3) but also to mark a goal role in the dative structure as in (4).

(3) *Otokonoko-ga* *tomodachi-ni* *tatak-are-ta.*
boy-NOM friend-OBLIQUE hit-PASSIVE-PAST
"The boy was hit by his friend."

(4) *Otokonoko-ga* *tomodachi-ni* *hanashikake-ta.*
boy-NOM friend-DAT talked to-PAST
"The boy talked to his friend."

This creates structural ambiguity on encountering a noun phrase with the case marker *ni* as comprehenders would not know whether it denotes an agent role in the passive structure or a goal role in the dative structure. We hypothesized that a dative use of the ambiguous case marker *ni* would increase the probability of predicting the dative on the next encounter with the case marker. In this study, we tested syntactic priming by presenting active/passive transitive prime sentences and seeing if they influenced descriptions of transitive target pictures. We also collected reading time measures to provide measures of expectations on the primes. To manipulate expectations about the case marker *ni*, we varied the fillers that were used between the prime-target pairs in three conditions. The All-dative condition only had dative fillers, the No-dative condition had other structures without *ni* arguments, and the Half-dative had half of both of these conditions.

One possibility is that participants will treat dative *ni* and passive *ni* as the same case marker based on their form. In this Form-based account, participants will have stronger expectations for the case marker *ni* in the All-dative condition than the No-dative condition (with Half-dative in between). If priming is due to error or surprisal, then they should show less syntactic priming of passives when there are more dative fillers. Thus, the Form-based account predicts priming will be strongest in the No-dative condition.

Another possibility is that participants treat dative ni and passive ni as different case markers based on their function. This Function-based account argues that the structures that are primed depend in part on the meaning that they encode. The case marker *ni* encodes the agent in the passive and it encode the recipient in the dative. If this is the case, then the All-dative condition will yield strong expectations that the case marker *ni* is marking the dative recipient, and when it is seen in the passive prime, strong error will be generated. Thus, the Function-based account should predict stronger passive priming in the All-dative condition than the No-dative condition.

A final possibility is motivated by theories of syntactic priming that do not use error/surprisal. In these theories, syntactic priming is due to the residual activation of structures that is caused by the processing of prime sentences (Pickering & Branigan, 1998). Since passive structures and dative structures are different structures, residual activation does not transfer from one structure to the other. In addition, the residual activation mechanism does not depend on expectations, so the composition of the fillers should not influence passive priming. Therefore, this activation account predicts priming should be the same across all three conditions.

2. Experiment

2.1. Method

2.1.1. Participants

Forty-three participants, recruited from the student community at the University of Tokyo, took part in the experiment for which they received monetary compensation. All the participants were native speakers of Japanese.

2.1.2. Stimuli

We prepared 18 experimental items, each consisting of a pair of a prime item and a target item. The prime item was either an active sentence (5a) or a passive sentence (5b).

(5a)	*Rooba-ga*	*akanboo-o*	*sarat-ta.*
	old lady-NOM	baby-ACC	abduct-PAST

“The old lady abducted the baby.”

(5b)	*Akanboo-ga*	*rooba-ni*	*saraw-are-ta.*
	baby-NOM	old lady-DAT	abduct-PASSIVE-PAST

“The baby was abducted by the old lady.”

The target item was made of a set of a verb and a target picture. The verb was presented prior to the target picture. The target picture was a simple line-drawing such as Figure 1.

The target picture can be described by either an active sentence (e.g., *Keikan-ga dorobou-o tsukamaeta*, "The policeman caught the thief") or a passive sentence (e.g., *Dorobou-ga keikan-ni tsukamaerareta*, "The thief was caught by the policeman"). The verb was never repeated between prime and target[2)]. We included sixty filler sentences in total. Two to four fillers were inserted between each experimental item. For those fillers, we manipulated the occurrence of the dative structure at three levels. In the All-dative condition, the fillers that appeared between experimental items were always dative sentences. In the Half-dative condition, half of the fillers were the dative structure and the other half were the unrelated structure. In the No-dative condition, all the fillers were of the unrelated structures. These unrelated fillers were neither the dative sentences nor the passive sentences. In addition, we

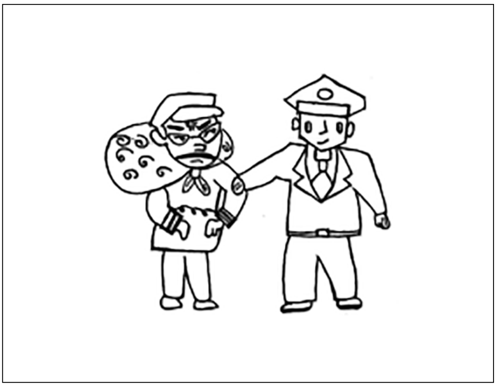

Figure 1. An example of target pictures

2) Previous research showed that repetition of the verb is not required for a priming effect of the passive structure in Japanese (Arai & Mazuka, 2014; See also Arai, 2012 for the discussion on the influence of head-directionality on syntactic priming).

also included 12 picture fillers for picture description, whose pictures depicted intransitive events. The independent variables are Prime Type (Active or Passive) and Filler Type (All-dative, Half-dative, or No-dative), resulting in 2 × 3 design. We created 6 experimental lists based on Latin square design.

2.1.3. Procedure

Participants were given an identical instruction and signed a consent form. They were then randomly assigned to one of these experimental conditions. They read and comprehend a sentence that was presented on the monitor and were asked to describe the picture when a picture was presented, using the verb presented prior to the picture. In the experimental trails, participants first read a prime sentence, which was either an active or passive sentence. They next saw a target verb (e.g., *tsukamaeru*, 'arrest') and then a target picture, which they described using the previously shown verb. Crucially, all the participants saw the same experimental items and only difference was the structure type of the filler sentences that are inserted between them. The entire experiment took approximately 30 minutes.

2.2. Results

2.2.1. Reading Times of Prime Sentences

We first analyzed reading times of the sentence-final verb in the prime sentences. Note that the verb form was different between the active and passive conditions but identical across the three filler conditions. Figure 2 shows the reading times at the verb in each condition.

For the data analysis, we first excluded reading times exceeding 5000ms as such extreme values are likely to reflect some kind of errors. We analyzed the remaining data using Linear Mixed-Effects (LME) models (e.g., Baayen, Davidson, & Bates, 2008). The model included Prime Type (Active or Passive) and Filler Group

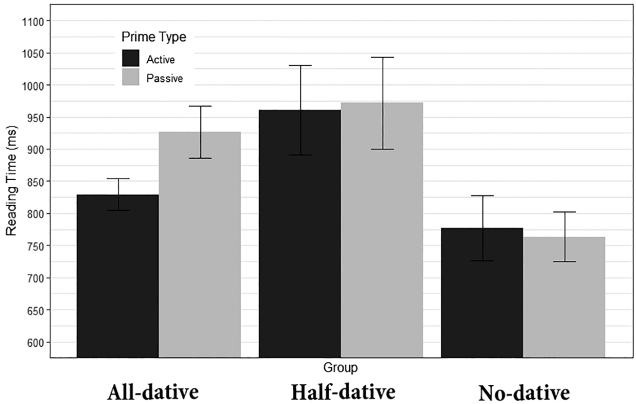

Figure 2. Reading times at the verb in prime trials for each condition

(All-dative, Half-dative, or No-dative) as fixed factors, including their interaction, plus Trial Order (The order number that the prime trial appeared in the experiment). Participants and items were included as random factors. Prime Type was centered with deviation coding (-0.5, + 0.5) while the filler structure was coded with dummy variable coding, with the No-dative condition is treated as a baseline. Trial Order was centered and normalized. We selected the optimal model using backward selection approach. We started from the model with the maximal random structure including slopes for Prime Type and Filler Group as well as their interaction for both random variables. With the optimal model, we further excluded extreme values whose residuals exceeded ±3 SD from the model (Baayen, 2008). We report coefficients (*β*), standard errors (SE), *t* values, and their *p* values from the best-fit model. Fixed factors' *p* values were computed based on the posterior distribution computed using Monte Carlo Markov chain sampling. Where the optimal model included more than one random slope, we computed the *p* value for each factor using likelihood ratio (LR) tests. Table 1 showed the results from the optimal model.

Table 1. The results from the optimal LME model on reading times at the verb

	Coeff.	*SE*	*t*	*p*
(Intercept)	754.5	77.5	9.74	<.001
Prime Type	14.1	35.5	0.40	0.692
Half-dative	173.3	107.7	1.61	0.115
All-dative	89.7	105.9	0.85	0.402
Trial Order	-80.1	17.53	-4.57	<.001
Prime Type × Half-dative	-9.4	50.3	-0.19	0.852
Prime Type × All-dative	99.0	49.5	2.00	0.046

The results showed that there was neither a main effect of Prime Type nor Filler Group. The main effect of Trial Order was fully significant, suggesting that the reading times of the verb decreased over the course of the experiment. Importantly, we observed a significant interaction between Prime structure and the All-dative condition. The positive coefficient suggests that the difference in reading times between active and passive prime sentences were significantly larger in the All-dative condition than in the No-dative condition. Given that our participants read the identical prime sentences, this must be due to the experience of the dative fillers. Since those in the All-dative condition expected the dative structure more after reading the dative fillers, they were more surprised to see the passive verb, resulting in increased reading times in the All-dative condition compared to those in the No-dative condition. The interaction between Prime Type and the Half-dative condition was not significant.

2.2.2. Production of passive descriptions

We next analyzed the production of passive descriptions. Figure 3 showed the proportions of the target descriptions using the passive structure.

We analyzed the data using Generalized Linear Mixed Models (GLMM) (Jaeger, 2008). We selected the optimal model using backward selection approach. The fixed and random effects included in the model was the same as the one on reading times (Prime Type as within-participants and Filler Group as between participants as well as Trial Order). Table 2 showed the results of the analysis.

There was neither a main effect of Prime Type nor that of Filler Group although the effect of the Half-dative condition was marginally significant. There was a main effect of Trial Order, suggesting that our participants became less likely to produce a passive description over the course of the experiment. This is likely to suggest that the participants become less surprised on seeing the passive prime over the course of the experiment. In the experiment as a whole, they were exposed to the equal production of the active and passive structures in the whole

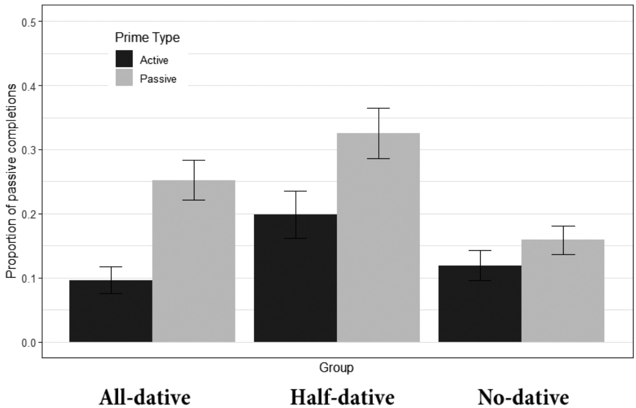

Figure 3. Proportions of passive completions per condition.

Table 2. The results from the optimal GLMM model on the number of passive descriptions

	Coeff.	*SE*	*z*	*p*
(Intercept)	-2.56	0.45	-5.68	<.001
Prime Type	0.60	0.41	1.45	0.147
Half-dative	0.99	0.55	1.81	0.070
All-dative	0.14	0.55	0.26	0.793
Trial Order	-0.74	0.22	-3.40	<.001
Prime Type × Half-dative	0.36	0.54	0.66	0.510
Prime Type × All-dative	1.16	0.58	2.00	0.046

experiment, which is different from normal statistics in real life, in which the passives occur relatively much less frequently than the actives. Importantly, we observed a significant interaction between Prime Type and All-dative condition. This demonstrates that the structural priming effect was larger in the All-dative condition compared to the No-dative condition. We did not find an interaction between Prime Type and Half-dative condition. The results of the passive production therefore are consistent with those of reading times at the verb.

2.2.3. Additional Analysis I: Combined Analysis of Reading Times and Passive Production

We next combined the reading time data in the prime trials and the passive production data in the target trials for each trial and conducted a combined analysis. If an effect of structural priming reflects prediction errors at the verb in the prime trials, we may be able to observe a direct relationship between the reading times and an effect of priming. We replaced Filler Group with the reading times in prime trials and used the same model as the one on the passive descriptions. We now have a continuous variable of Prime Reading Time instead of the three-level categorical factor of Filler Group. Table 3 showed the results of the analysis.

Figure 4 shows the relationship between the reading times at the verb in the prime trials and the likelihood of the passive production for each prime structure.

Table 3. The results from the optimal GLMM model of the combined analysis

	Coeff.	*SE*	*z*	*p*
(Intercept)	-2.16	0.31	-7.02	<.001
Prime Type	1.09	0.24	4.60	<.001
Prime Reading Time	0.12	0.14	0.83	0.405
Trial Order	-0.75	0.22	-3.49	<.001
Prime Type × Prime Reading Time	-0.48	0.22	-2.20	0.028

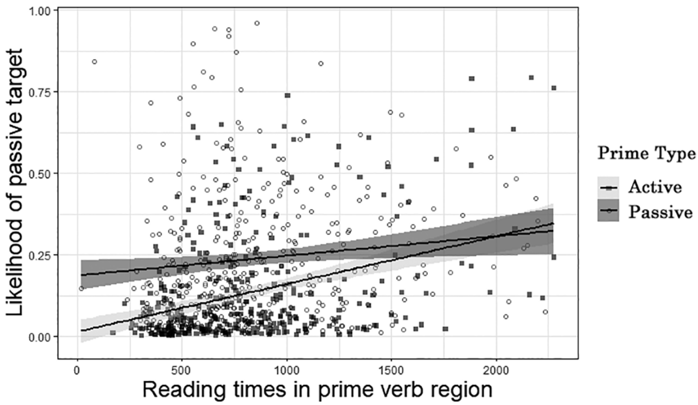

Figure 4. Relationship between predicted likelihood of passive target and reading times in verb region of prime sentences

Although there is a numerical trend that the greater reading times are associated with the greater likelihood of the passive production, the main effect of Prime Reading Times was not significant. However, the results showed a significant interaction between Prime Type and Prime Reading Time.

The pattern of the interaction however does not exactly match what we predicted. It appears to reflect the positive correlation between the likelihood of the passive production and the reading times of the active prime sentences rather than those of the passive prime sentences. Planned comparisons demonstrated that the simple effect of Prime Reading Times is not significant with passive primes (β = -0.12, SE = 0.16, z = -0.76, p = 0.445) and it is marginally significant with active primes (β = 0.35, SE = 0.19, z = 1.86, p = 0.062). Although the effect was not fully significant, this indicates that the greater reading times at the verb in active prime sentences were associated with the greater likelihood of the passive production. We discuss possible interpretation of this finding in the following section.

2.2.4. Additional Analysis II: Interaction between Prime Structure and Trial order

In the above analyses, trial order was included as a control variable without including an interaction between Prime Type and Trial Order. Since the effect of word-level prediction errors should vary on a trial-by-trial basis due to adaptation to the input in the local environment, the influence of trial order on priming of the passive structure is non-trivial for the model of error-based implicit learning as well as the surprisal theory. We initially did not include this interaction in the earlier analyses because the models with the interaction failed to converge and thus yielded no output, which might be because there were not enough data points in each condition in this complex design.

One different approach is to adopt Helmert coding, which can be applied for ordinal variables. With this coding scheme, each level of a variable is compared to *later* levels of the variable. Since the Filler Group variable can be treated as an ordinal variable (All→Half→None), this coding scheme could solve the convergence issue with the earlier analyses. Specifically, two contrasts are tested with this coding. The first one is the contrast of No-dative level vs. two dative levels (All-dative and No-dative conditions merged). The second one is the contrast of Half-dative level vs. All-dative level (excluding the No-dative level). Inclusion of the interaction between Prime Type and Filler Group with Helmert coding would tell us how the differences in these contrasts would differ for each prime structure. This additional analysis with Helmert coding also included a three-way interaction between Prime Type, Filler Group, and Trial Order both on the prime reading times as well as the target passive descriptions. First, Table 4 shows the output from the optimal LME analysis on the reading times at the verb in the prime trials.

The results showed a significant interaction between Trial Order and the Datives vs. No-dative contrast and also a significant interaction between Trial Order and Prime Type. Importantly, these two interactions are accounted for by a significant three-way interaction between the Datives vs. No dative contrast, Prime Type, and Trial Order. Figure 5 shows the relationship between the prime reading times and trial order for each group and prime type. This three-way interaction is due to the fact that the lines for actives and passives differ in the two dative filler

Table 4. The results from the optimal LME model on reading times with the interactions with trial order included

	Coeff.	*SE*	*z*	*p*
(Intercept)	6.63	0.05	130.76	< 0.001
Trial Order	-0.10	0.02	-6.05	< 0.001
Prime Type	0.05	0.02	2.43	0.016
Datives vs. No-dative	0.11	0.07	1.58	0.121
All-dative vs. Half-dative	-0.03	0.12	-0.25	0.807
Datives vs. No-dative × Trial Order	0.04	0.03	2.29	0.023
Prime Type × Trial Order	-0.05	0.02	-2.52	0.012
Prime Type × Datives vs. No-dative	0.02	0.03	0.77	0.444
All-dative vs. Half-dative × Trial Order	-0.03	0.03	-1.17	0.242
All-dative vs. Half-dative × Prime Type	0.12	0.05	2.34	0.020
Datives vs. No-dative × Prime Type × Trial Order	-0.06	0.03	-2.01	0.045
All-dative vs. Half-dative × Prime Type × Trial Order	0.02	0.05	0.38	0.706

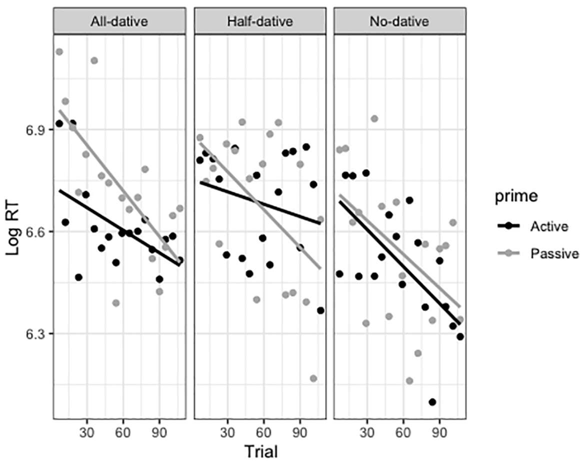

Figure 5. Relationship between log-transformed reading times of prime verbs and trial order for each prime type

conditions, with the slope for the passive prime being steeper than that for the active prime, but are almost parallel in the No-dative condition.

This demonstrated the longer reading times for the passive structure at the beginning of the experiment compared to the active structure in the two dative conditions, suggesting that the impact of dative fillers on the prediction error with passive prime sentences was largest at the beginning of the experiment.

Next, Figure 6 shows the relationship between the production of the passive structure and trial order for each group and each prime type. The pattern looked surprisingly similar to that of the reading times. Whereas there seems no difference between the two prime types in the No-dative condition, the likelihood of passive production in the passive prime condition seems higher at the beginning of the experiment than that in the active prime condition in the two dative filler conditions. However, the GLMM with Helmert coding failed to converge and no output was obtained. The pattern was therefore not statistically confirmed.

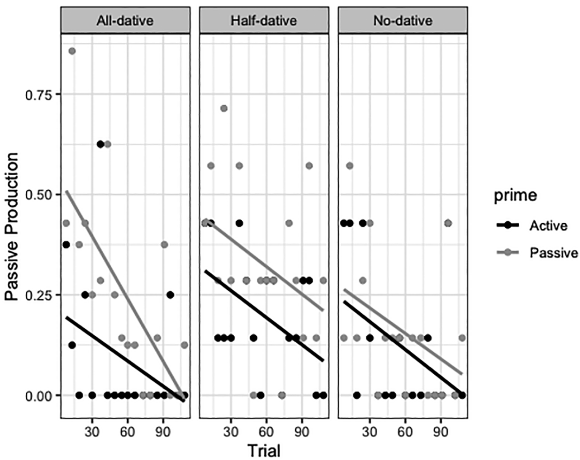

Figure 6. Relationship between the likelihood of passive production and trial order for each prime type

In sum, the additional analysis of reading times showed that the processing cost that reflects prediction errors on the passive prime sentences following dative fillers was largest at the beginning of the experiment and gradually decreased over the course of the experiment. Consistently, although not statistically confirmed, the pattern of the results for the production of passive descriptions was consistent with that of the prime reading times.

3. Discussion

The current study investigated an influence of word-based predictions on structural priming. We manipulated the structure of filler sentences inserted between experimental items. Taking advantage of the ambiguity of the case marker *ni*, we created an environment in which participants were led to expect a different structure on encountering the noun phrase case-marked with *ni* in prime sentences.

The results of the reading times at the verb in the prime sentences showed that the prior experience of dative fillers in the All-dative condition increased the reading times at the verb in the passive prime sentences, suggesting that dative fillers led to greater expectations of the dative structure and thus resulted in greater predictor errors on encountering the verb in the passive form. Importantly, prime-target pairs were identical across all the filler group conditions, ensuring that the effect was due to the processing of filler sentences, particularly, that of the dative structure with the ambiguous case marker *ni*.

The results from the production data showed that the size of structural priming was larger when participants experienced the dative structure in filler items compared to when they did not experience the dative structure at all. This finding suggests that the experience with the dative fillers, which caused larger prediction error on encountering the passive structure in prime trials, resulted in greater syntactic priming. Our results thus provided evidence for the Function-based account, according to which prediction error is sensitive to the function of the case marker *ni* and thematic role information. Specifically, after reading dative fillers, comprehenders were more likely to predict another dative sentence on encountering a noun phrase with the case marker *ni* following the nominative argument in prime sentences. This caused greater surprisal or prediction error on encountering the subsequent verb in passive form, which in turn resulted in large weight changes in the networks representing the active and the passive structures. This was reflected in the priming effect of the passives. Our results provide some of the first evidence for the word-level prediction errors on structural priming.

In addition, the combined analysis of the reading times of prime sentences and the production of the passive descriptions in target trials showed a significant interaction between Prime Type and Prime Reading Time. However, the results were not the predicted pattern and were likely due to the positive correlation between the likelihood of the passive production and the reading times of the active prime sentences rather than those of the passive prime sentences. This prima facie appears rather odd as the active prime sentences are not supposed to cause word-level prediction errors. One possibility is that in some prime trials of the active structure comprehenders may have predicted the passive structure before reading prime sentences. This sentence-level prediction was disconfirmed by an active prime sentence, resulting in increased reading times at the verb. This sentence-level prediction error could have led to the greater likelihood of the passive production. In fact, the prediction of the passive structure becomes quite reasonable after encountering a certain number of experimental trials since the frequency of passive sentences relative to that of active sentences in the experiment (i.e., 1:1) is so much higher than that in real life. Although this is only a possibility, this may suggest that both word-level prediction errors and the sentence-level prediction errors contribute to the priming effect of the passive structure in Japanese.

Furthermore, another set of additional analyses was conducted to examine an influence of trial order on the effects of prime structure and filler group. The results showed that participants initially found the active easier to process than the passive (Prime Type), but they came to expect both structures more as the experiment progressed (Prime Type × Trial Order). But this pattern depended on the fillers in the study as when the fillers had no datives, participants quickly came to expect both prime structures and they processed them at the same speed (Datives vs. No-dative × Prime Type × Trial Order). In contrast, in the conditions with dative fillers, participants processed active and passives differently at different points in the study, suggesting that the dative *ni*-arguments made it harder to predict the meaning of the agent *ni*-argument in the passive prime structure. Although not fully supported by statistical analyses, the similarity of the pattern for the production data suggests that greater difficulty in processing (prediction error) was positively related to greater priming in production, as suggested in error-based learning theories.

These results are not compatible with the activation-based model. The activation-based model assumes that residual activation is left in nodes for the active and passive structures and this residual activation causes syntactic priming in the transitive alternation. But the dative structure involves different nodes, so there is no reason to expect any transfer of residual activation to active and passive transitive nodes. Previous work has found priming from structures with transitive verbs to structures with dative verbs (Chang, Tsumura, Minemi, & Hirose, 2022), but this only occurred when the transitive verb occurred with three arguments (*ga*/*ha* + *ni* + *o*). In a node-based account, it would be argued that three-argument sentences used shared structural nodes and this explains why priming transfers from transitives to datives. But in both the Form-based and Function-based accounts, we assume that syntactic representations are more distributed with separate elements for particles and structures, and these sentence elements may be linked to distinct thematic roles. Therefore, the previous work that found transitive to dative priming in Japanese is consistent with the present work where dative *ni* arguments affected the processing of transitive *ni* arguments.

This study presented dative filler sentences and the expectations about these structures presumably increased over the study. This was, however, not reflected in reading times of passive sentences, which decreased over the study due to a large number of experiences of the passive structure. This issue is related to the phenomenon called cumulative priming, where priming of a particular structure builds up over the experiment. Several studies have manipulated the participant's recent experience with particular structures. For example, Kaschak and colleagues have varied the number of DO and PO structures that precede a priming phase, and varied whether they are blocked or not (Kaschak, 2007; Kaschak & Borreggine, 2008). Some of these studies have not found strong variation in priming due to frequency or blocking (e.g., Kaschak, Loney, & Borreggine, 2006). Jaeger and Snider (2013) did a reanalysis of several of Kaschak's studies and found that structural surprisal did predict greater priming when the probability of a structure was changed as each sentence was processed. Although these studies provide a mixed set of results about the role of cumulative priming and error-based prediction, one thing to remember is that our study used different structures for fillers and prime trials. This study therefore provides some of the first evidence that cumulative priming for one structure (e.g. datives) can influence priming of a different structure (e.g. transitives).

In summary, being consistent with the error-based model, our results demonstrated that syntactic priming of the passive structure was caused by the errors in word-level predictions during the incremental processing of the prime sentences. Furthermore, the influence of word-level prediction errors was observed in the environment in which word-level predictions was not dependent on particular lexical items. The effect was observed without the repetition of the verb between prime and target, ensuring that the errors are not lexically associated. This is important because our findings are independent of an influence of explicit memory. The current study thus provided the first evidence for the influence of word-based prediction errors on lexically independent syntactic priming.

Acknowledgments

We thank Megumi Ishikawa for helping to conduct the experiment and collect data reported in the current paper.

4. References

Arai, M. (2012). What can head-final languages tell us about syntactic priming (and vice versa)? *Language and Linguistics Compass*. *6*, 545-559.

Arai, M., & Mazuka, R. (2014). The development of Japanese passive syntax as indexed by structural priming in comprehension. *The Quarterly Journal of Experimental Psychology*, *67*, 60-78.

Arai, M., van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *54*, 218-250.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*, 455-461.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-362). New York: Wiley.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355-387.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology-General*, *129*, 177-192.

Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 468-481.

Branigan, H. P., Pickering, M. J., Stewart, A. J. & McLean, J. F. (2000). Syntactic priming in spoken production: Linguistic and temporal interference. *Memory & Cognition*, *28*, 1297-1302.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*. 234-72.

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*, 217-229.

Chang, F., Tsumura, S., Minemi, I., & Hirose, Y. (2022). Abstract structures and meaning in Japanese dative structural priming. *Applied Psycholinguistics*, *43*, 411-433.

Fine A. B., Jaeger T. F., Farmer T. A., & Qian T. (2013). Rapid Expectation Adaptation during Syntactic Comprehension. *PLoS ONE*, *8*: DOI: 10. 1371/journal. pone. 0077661.

Fine, A. B., & Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*. *37*, 578--591.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics* (pp. 159-166). Pittsburgh, USA: Association for Computational Linguistics.

Hartsuiker, R. J., Bernales, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*, 214-238.

Ishikawa, M. Arai, M. & Hirose, Y. (2014). Priming of the passive structure from short and full passives with Japanese children and adults. *Information and Communication Engineers IEICE Technical Report*, *114*, 133-138.

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of memory and language*, *59*, 434-

446.

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*, 57-83.

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*, 133-156.

Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition*, *35*, 925-937.

Kaschak, M. P., & Borreggine, K. L. (2008). Is long-term structural priming affected by patterns of experience with individual verbs? *Journal of Memory and Language*, *58*, 862-878.

Kaschak, M. P., & Glenberg, A. M. (2004). This construction needs learned. *Journal of Experimental Psychology: General*, *133*, 450-467.

Kaschak, M. P., Loney, R. A., & Borreggine, K. L. (2006). Recent experience affects the strength of structural priming. *Cognition*, *99*. 73-82.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126-1177.

Miyamoto, E. T. (2002). Case markers as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research*, *31*, 307-347.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*. *39*, 633-651.

Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*, *58*, 188-213.