Syntactic cues may not aid human parsers efficiently in predicting Japanese passives

Masataka Ogawa

Proceedings of the 38th Pacific Asia Conference on
Language, Information and Computation (PACLIC 38)

Nathaniel Oco, Shirley N. Dita, Ariane Macalinga Borlongan, Jong-Bok Kim (eds.)

2024

Syntactic cues may not aid human parsers efficiently in predicting Japanese passives

Masataka Ogawa
The University of Tokyo / 3-8-1, Komaba, Meguro-ku, Tokyo, 153-8902, Japan
ogawa.phiz@gmail.com

Abstract

Our study demonstrates that the human parser may not predict passive constructions from syntactic elements preceding the sentence-final verb in Japanese by comparing the reading time and comprehension accuracy of V-(*r*)*are* passive and V-∅ active sentences. In SVO languages like English, where the syntactic structures of actives and passives differ, reading times for passives are often shorter, and comprehension accuracy is comparable for both constructions. However, in Japanese, an SOV language, where the syntactic structures of actives and passives are similar, prior studies found numerically longer reading times and lower comprehension accuracy for passives. We hypothesized that if reading times for passives were shorter as in SVO languages, a case marker in passives might signal the passive construction and reduce reading times for passives. Controlling verb classes that assign different case markers to non-subject NPs, we carried out a self-paced reading (SPR) task where participants read sentences at their own pace, to determine if syntactic cues facilitate the prediction of V-(*r*)*are* before the sentence-final verb. A comprehension question to assess comprehension accuracy followed each trial of the SPR task. The results did not reveal that differences in case markers led to faster reading times or higher accuracy for passives. Rather, we corroborated the previous findings: increased reading times and lower accuracy for Japanese passives.

There are contradicting views on how reading times and comprehension accuracy are different between passives and actives in SVO and SOV languages. Studies on SVO languages suggest that the processing load for passives is the same or less than for actives, with parsers predicting passives as they read (e.g. Paolazzi et al., 2016, 2017, 2019). Conversely, research on Japanese, an SOV language, have indicated that passives lead to processing difficulties (Tamaoka et al., 2005; Kinno et al., 2008; Tanaka et al., 2017). Even in experiments with equivalent morphological complexity of both active and passive verbs, reading times for passives were longer, and comprehension was less accurate (Ogawa, 2023). However, these previous research did not clarify if a passives can be predicted from syntactic cues before the sentence-final verb.

We performed a self-paced reading (SPR) task, where participants read sentences at their own pace, to determine if syntactic cues facilitate the prediction of Japanese passives V-(*r*)*are* before the sentence-final verb. While comparing reading times between the passive and its active counterpart V-∅, we controlled verb classes assigning different case markers to non-subject NPs, hypothesizing that certain markers predict passives. However, we found no evidence that the human parser predicts passive construction from the case marker in Japanese. Instead, we replicated robust findings of longer reading times and lower accuracy for passives using V-(*r*)*are* and V-∅, which were not employed in a previous SPR experiment (Ogawa, 2023).

Section 1 reviews contradicting results in various languages on reading time of passives, and explain why Japanese case marker can contribute to the prediction of voice/diathesis. Section 2, their comprehension accuracy. Section 3 outlines the methodology and Section Section 4 reviews results of the experiment.

1 Can passives be read faster?

Paolazzi et al. (2016; 2017; 2019) performed SPR experiments in English and discovered shorter or equivalent reading times for verbs and post-verbs

The glossing abbreviations in this article follow Leipzig Glossing Rules (Department of Linguistics of Max Planck Institute for Evolutionary Anthropology, 2008, last accessed on July 15, 2022), Brown and Anderson (2006), , except INFR. -: affix boundary / =: clitic boundary / ACC: accusative / ADV: adverb / DAT: dative / INFR: inferential mood / NOM: nominative / PASS: passive / POL: polite register / PST: past / Q: question particle in passive sentences, compared to those in active sentences. They suggested that the auxiliary verb *be* and the preposition *by* in passive constructions aid in predicting the (post-)verb region, as the auxiliary verb *be* signals the upcoming presence of a verbal past participle. Paolazzi et al. (2019; 2021b) argue that in passives, elements before the verb increase predictability for verbs, leading to reduced reading times. They noted that, as the verb and preposition *by* signal a subsequent non-subject NP in passive, such a NP is more predictable in passives. This increased predictability reduces reading times for the post-verbal region in passives, in contrast to actives where only the verb serves as a cue for the region.

However, in Japanese, reading times were numerically longer for passive verbs, although there was unclear statistical support (Ogawa, 2023). They compared reading times using benefactive V-*te morau* passives and V-*te ageru* actives, where the morphological complexity of the verbs in both constructions was equalised. This suggests that the delay in reading times is likely due to the process of associating thematic roles with grammatical relations in passive sentences rather than differences in morphological structure between active and passive verbs. Furthermore, Ogawa (2023) argued that the difference in reading times between active and passive sentences in previous studies on English is due to the fact that only passives have morphosyntactic cues in English.

Indeed, it is challenging to control for morphosyntactic complexity when comparing reading times between actives and passives in English. Paolazzi et al.'s (2019) SPR experiment compared reading times for the past-tensed main verb in active sentences and the past participle in passive sentences as the same region. Both constructions contained a subject NP preceding the verb, but only passives include the copula *be* in an additional region. This created an imbalanced design where only the passives had a predictor (*be*) for the passive voice, whereas actives lacked any corresponding predictor. However, this issue can be avoided by using SOV languages like Japanese.

As the verb appears at the end of the sentence in Japanese, markers that signal the sentential diathesis would necessarily precede the verb, if such exist. Moreover, the structure of the subject and the object/oblique NP can be very similar in Japanese, with the only difference being the case marker (adposition) attached to the object/oblique NP, as shown in (1).

(1) a. V-(*r*)*are* passive; =o_{ACC}-verb
Takahashi=ga *Ōtsuka=ni*
T.=NOM Ō.=DAT
naguritobas-**are**-*ta.*
hit-PASS-PST
'Takahashi was punched by Ōtsuka.'

b. V-∅ active; =o_{ACC}-verb
Ōtsuka=ga *Takahashi*=**o**
Ō.=NOM T.=ACC
naguritobashi-ta.
hit-PST
'Ōtsuka punched Takahashi.'

Since Japanese adpositions are consistently present in both active and passive sentences, this avoids the imbalance of having adpositions in one construction but not the other, and allows for a clearer comparison to examine whether the human parser predicts a passive sentence when reading the adposition attached to the oblique NP, if the predictors of the sentence diathesis are adpositions. In fact, it is plausible that the ease of predicting actives versus passives in Japanese varies depending on the adposition used.

Muraoka (2006) had participants complete sentences by filling in a sentence-final VP after being presented with subject and non-subject NPs. Results indicated that predictions for what follows the non-subject NP depend on its case marking (see also Figure 3 in Appendix A.). Muraoka (2006) suggested that a =ni_{DAT}-marked NP predicts either an =o_{ACC}-marked NP (forming a ditransitive construction) or a verb, while an =o_{ACC}-marked NP predicts a verb will directly follow.

Muraoka (2006) did not specify which voice is predicted when encountering a =ni_{DAT}-marked NP or =o_{ACC}-marked NP. However, their data indicate that passive verbs were predicted with a =ni_{DAT}-marked NP, but not with an =o_{ACC}-marked NP. Hence, only a =ni_{DAT}-marked NP, not an =o_{ACC}-marked NP, could signal that the parser is reading a passive sentence. If so, the reading time difference introduced by such a voice/diathesis prediction can be found between the =ni_{DAT}-marked and =o_{ACC}-marked NP.

Moreover, in Japanese active, the accusative =*o* marks the object for some verbs (=o_{ACC} verbs, (1)), whereas the dative =*ni* marks the object for others (=*ni*$_{\text{DAT}}$ verbs, (2)). We can utilise this asymmetrical case pattern to test whether Japanese case marker can signal the voice of subsequent VP or the diathesis of entire sentence.

(2) a. V-(*r*)*are* passive; =*ni*$_{\text{DAT}}$-verb

Takahashi=ga	*Ōtsuka=ni*
T.=NOM	Ō.=DAT

nagurikakar-**are**-*ta.*
hit-PASS-PST
'Takahashi was lunged at by Ōtsuka.'

b. V-∅ active; =*ni*$_{\text{DAT}}$-verb

Ōtsuka=ga	*Takahashi*=**ni**
Ō.=NOM	T.=DAT

nagurikakat-ta.
hit-PST
'Ōtsuka lunged at Takahashi.'

2 Are passives comprehensible?

Paolazzi et al. (2021b) noted that processing difficulties for passives in English arise during comprehension questions written in active voice inquiring thematic roles, such as questions asking *who* performed an action on *whom*. They showed that participants responded less accurately to active voice comprehension questions about thematic relations of passive target sentences. Similar findings were also reported in German (Grillo et al., 2019; Meng and Bader, 2020).

In contrast to the findings in SVO languages, several studies in Japanese have indicated that the passive constructions using V-(*r*)*are* impose greater processing difficulties compared to their active counterparts (Tamaoka et al., 2005; Yokoyama et al., 2006; Kinno et al., 2008; Tanaka et al., 2017). Tamaoka et al. (2005), for instance, carried out experiments in which participants judged the sensibility of various sentence structures, including active and passive constructions presented in the canonical SO order and non-canonical OS order of NPs. Longer reaction times were found for passives than for actives in both word order conditions, despite nearly equivalent error rates. These results suggested that human parsers encounter a larger processing cost when comprehending passives.

Several fMRI studies found that when participants judged whether a written V-(*r*)*are* passive correctly described a picture of one stick figure acting on another, more activation was triggered in the left inferior frontal gyrus compared to the corresponding active sentences (Kinno et al., 2008; Tanaka et al., 2017). Kinno et al. (2008) concluded that the syntactic reanalysis occurred to comprehend the patient denoted by a =*ga*-marked nominative NP in passives. However, Yokoyama et al. (2006) observed a similar activation in that cerebral region, when they compared the cognitive demands of uninflected V-∅ active verbs and inflected V-(*r*)*are* passive verbs in a lexical decision task. They concluded that unmarked active verbs are treated as unitary words, while marked passive verbs involve morphological decomposition. Thus, a brain activity specific to passives is expected, although it is arguable whether this is caused by processing diathesis (entire sentence level) or voice (verbal morphological level).

Further evidence for lower comprehension accuracy of Japanese passives comes from Ogawa (2023), which employed a similar comprehension question paradigm as Paolazzi et al. (2021b). They minimized the morphological difference between active and passive sentences, which was a limitation of previous studies, by using benefactive active/passive pairs (V-*te ageru* and V-*te morau*). Therefore, they concluded that the observed decrease in accuracy for passive comprehension was caused by the cognitive process that links the patient to the grammatical subject in passives, rather than morphological factors.

3 Self-paced reading experiment with comprehension question

Existing literature has provided evidence that passive sentences in Japanese demand more time to read and present greater difficulties for precise understanding. Nevertheless, the potential role of the dative case marker =*ni*$_{\text{DAT}}$ in passives as a signal for the passive construction, which could consequently decrease reading times, has not been extensively explored. Thus we explored two key issues: first, we investigated whether the parser predicts a passive voice for the subsequent VP upon reading a =*ni*$_{\text{DAT}}$-marked NP in the oblique region, thereby initiating constructing a passive structure at this or the post-oblique region. If so, the processing load for constructing the passive structure would increase reading times in the pre-verbal region (i.e. a =*ni*$_{\text{DAT}}$-marked NP) under the passive condition compared to the active condition. Second, we examined whether passives incur a greater parsing cost compared to actives at the verb and later regions.

To achieve these objectives, we employed an SPR experiment using a moving window paradigm (Just et al., 1982). We also appended a comprehension question task after each trial of the SPR experiment. This was to assess if Japanese V-(*r*)*are* passives, relative to V-∅ actives, impose a higher processing load to comprehend.

3.1 Participants

The same participants who were recruited for a previous study (Ogawa, 2023) also participated in the current experiment. Full details can be found in that paper. Note, however, that a total of 262 native Japanese speakers were recruited online, and we excluded eight participants from our analyses who did not meet the native speaker criteria.

3.2 Stimuli

3.2.1 Target sentences

As outlined in Table 1, we controlled the voice by employing V-∅ active or V-(*r*)*are* passive as the main verb chunk (R5). We also manipulated the oblique marker in R3 by using $=o_{\text{ACC}}$ and $=ni_{\text{DAT}}$. If $=ni_{\text{DAT}}$ in Japanese functions similarly to the passive predictors *be* and *by* in English (Paolazzi et al., 2019, 2021b), it would signal the human parser that the entire sentence is passive. Consequently, reading time would increase only for active sentences with $=ni_{\text{DAT}}$-verbs. This increase occurs because the parser, predicting a passive sentence after encountering $=ni_{\text{DAT}}$ in R3, experiences a surprisal effect when discovering that the sentence is actually active in R5.

This required the verb class in R5 to be a verb that take a $=ni_{\text{DAT}}$-marked object ($=ni_{\text{DAT}}$-verb) or those that take a $=o_{\text{ACC}}$-marked object ($=o_{\text{ACC}}$-verb). $=ni_{\text{DAT}}$-verbs are much rarer than $=o_{\text{ACC}}$-verbs. However, a number of verbal compounds consisting of two verbs (V-V compounds) take a $=ni_{\text{DAT}}$-marked object, while others take a $=o_{\text{ACC}}$-marked object. These $=ni_{\text{DAT}}$- and $=o_{\text{ACC}}$-V-V compounds were selected from the lexical compound verbs listed in the Compound Verb Lexicon (Kageyama, 2013). These lexical compounds are assumed to be registered in the lexicon due to their strong unity as words, preventing other grammatical elements from being inserted between the two verbs. It is unlikely that such V-V compounds are derived by syntactic operations (Kageyama, 1993). We also confirmed that both lexical $=ni_{\text{DAT}}$- and $=o_{\text{ACC}}$-V-V compounds chosen for target sentences can be used as passive verbs to a similar extent, based on the high MI and LogDice scores reported in NINJAL-LWP for BCCWJ (National Institute for Japanese Language and Linguistics and Lago Institute of Language, 2012).

We employed verbs corresponding to Type 1 'Direct effect on patient' in the hierarchy of two-place predicates proposed by Tsunoda (1985; 2009), to use eventive passive sentences for the passive condition as in previous studies of English and German (Paolazzi et al., 2016, 2017, 2019, 2021a,b; Grillo et al., 2019; Meng and Bader, 2020).

In line with earlier research (Witzel and Witzel, 2011; Koizumi and Imamura, 2017; Ogawa, 2023), we measured reading times in the verb region (R5) and the following modal particle region (R6) as indicators of cognitive load during the processing of verbal voice and sentential diathesis. The load elicited in R5 may spill over to R6 (Just et al., 1982, 232–233) or manifest later, prolonging reading times in R6 (delay, Just et al., 1982, 236). Thus, increased reading time could potentially occur in R5, R6, or both. Analogous to the inclusion of R6, we placed an action-denoting adverb (R4) after the oblique NP (R3). This design allowed us to detect any cognitive load related to the prediction of a passive structure triggered by the oblique NP before reading the verb.

3.2.2 Questions to measure comprehension accuracy

Each V-(*r*)*are* passive and V-∅ active target sentence in the SPR tasks was paired with a variant of the questions exemplified in Appendix B. These questions aimed to test whether participants correctly interpreted the thematic relation of each target. These questions were derived from the first NP (NP1; R2), second NP (NP2; R3), verb (R5), and modal (R6) of the target sentences. We counterbalanced the correct responses ("yes" or "no") by presenting NP1 and NP2 in the questions in either the same sequence as in the trials of SPR task or in the inverse order.

To investigate the potential facilitatory effect of voice priming between a question and its target, as observed by Ogawa (2023) for Japanese benefactive active and passive sentences, we also counterbalanced the voice of the target sentences and comprehension questions. This resulted in two conditions: (1) a matched condition, in which an active question was paired with an active target, and a passive question with a passive target; and (2) a mismatched condition, in which an active question was paired with a passive target or vice versa.

Voice	Verb class	R1: Locative ADVP	R2: First NP [NP1]	R3: Second NP [NP2]	R4: ADV on action	R5: Verb	R6: Modal particle
active	=o_{ACC}-verb	*Kyōshitsu=de*	*Takahashi=ga*	***Ōtsuka=o***	*chikarazuyoku*	***naguritobashi-ta***	*rashī*
		classroom=LOC	T.=NOM	Ō.=ACC	forcefully	hit-PST	INFR
		'In the classroom, Takahashi seems to have forcefully punched Ōtsuka.'					
	=ni_{DAT}-verb	*Kyōshitsu=de*	*Takahashi=ga*	***Ōtsuka=ni***	*chikarazuyoku*	***nagurikakat-ta***	*rashī*
		classroom=LOC	T.=NOM	Ō.=DAT	forcefully	hit-PST	INFR
		'In the classroom, Takahashi seems to have lunged at Ōtsuka with a powerful punch.'					
passive	=o_{ACC}-verb	*Kyōshitsu=de*	*Takahashi=ga*	***Ōtsuka=ni***	*chikarazuyoku*	***naguritobas-are-ta***	*rashī*
		classroom=LOC	T.=NOM	Ō.=DAT	forcefully	hit-PASS-PST	INFR
		'In the classroom, Takahashi seems to have forcefully been punched by Ōtsuka.'					
	=ni_{DAT}-verb	*Kyōshitsu=de*	*Takahashi=ga*	***Ōtsuka=ni***	*chikarazuyoku*	***nagurikakar-are-ta***	*rashī*
		classroom=LOC	T.=NOM	Ō.=DAT	forcefully	hit-PASS-PST	INFR
		'In the classroom, Takahashi seems to have been lunged at Ōtsuka with a powerful punch.'					

Table 1: Experimental conditions with a sample item for the SPR task

We confirmed the grammaticality of all stimuli, including 16 target and 48 distractor sentences in the main trials and six practice items.

3.3 Procedure

We employed PennController for Internet Based Experiments (PCIbex; https://farm.pcibex.net/), a web application for psycholinguistic research. Participants accessed the site solely from their personal computers, and access from any mobile device was restricted.

A video introduction outlining the experimental design was automatically shown to participants. The video clarified that each of the 64 trials would involve an SPR task followed by a comprehension question. Participants completed six practice trials preceding the main experiment to familiarise themselves with the protocol.

In the SPR task, stimuli were initially masked by underscores, with each region unveiled sequentially upon pressing the space bar. Sentences were presented without inter-word or inter-region spaces, adhering to the standard Japanese typesetting. The stimuli were displayed using the Noto Sans Japanese font in black on a white background.

Upon completing the last region of a sentence, participants pressed the space bar to trigger a comprehension question, which was fully displayed immediately. Participants answered by selecting either the F key to indicate 'yes' or the J key for 'no'. The experiment withheld feedback on the accuracy of the answers. The correct answers ('yes' or 'no') were counterbalanced across targets and distractors during the experiment.

Following each question, a prompt instructed participants to press the space bar when ready to start the next trial. This message remained on the screen until the participant chose to proceed, allowing them to control the pace of the experiment.

The aforementioned procedure follows the method outlined in Ogawa (2023). However, this experiment uniquely counterbalanced several factors unlike previous studies: the voice of the target sentence (i.e., V-∅ active versus V-(*r*)*are* passive), the verb class (i.e., =ni_{DAT}-verbs versus =o_{ACC}-verbs), the voice of the comprehension question (i.e., V-∅ active versus V-(*r*)*are* passive), and the correct responses (i.e., whether 'yes' or 'no' was correct). Thus, one of 16 stimulus lists was presented following a Latin-square design.

3.4 Data exclusion criteria

We excluded data from 55 participants who either participated multiple times or were suspected of doing so. Data from 50 participants were also discarded due to improper presentation of stimuli or suspicion thereof. Moreover, data from two participants were removed because of recording errors on the server. Adopting Paolazzi et al's (2019) criterion, we excluded data from four participants whose overall accuracy for distractors was below 75%. Consequently, the final analysis included data from 143 participants.

For the analysis of reading time data, we excluded trials where participants incorrectly answered the corresponding comprehension question. We further filtered out reading times less than 80 ms from the data, following Paape et al. (2021), as this duration is considered the minimum time required for linguistic information to affect oculomotor control (Altmann, 2011).

3.5 Statistical analyses

We fit Bayesian generalised linear mixed models using the `brms` package (Burkner, 2021) in `R` (R Core Team, 2021). The models included correlated varying intercepts and slopes for participants and items. In brms, cmdstanr (Gabry and Češnovar, 2021) estimated coefficients and bridgesampling (Gronau and Singmann, 2021) computed Bayes factors based on stanfit objects transferred rstan (Guo et al., 2021). Models were run with four chains and 2,000 warm-up and 50,000 post-warm-up iterations in each chain. The NUTS sampler was configured to target a mean acceptance probability $\delta = 0.9$.

We evaluated the impact of each explanatory variable on the response variables (reading time and accuracy) by calculating Bayes factors BF_{10}. They provide the quantitative support for the alternative model, which incorporates the explanatory variable of interest, in comparison to the null model lacking that variable. A $\mathrm{BF}_{10} > 1$ indicates that the explanatory variable has an effect on the response variable, whereas a $\mathrm{BF}_{10} < 1$ indicates the absence of an effect. We adopted Lee and Wagenmakers's criteria (2013, derived from Jeffreys, 1939/1998) to interpret the strength of evidence for the presence or absence of an effect, as shown in Table 2.

BF_{10}	Strength of evidence
For the alternative model	
$100 < \mathrm{BF}_{10}$	Extreme
$30 < \mathrm{BF}_{10} \leq 100$	Very strong
$10 < \mathrm{BF}_{10} \leq 30$	Strong
$3 < \mathrm{BF}_{10} \leq 10$	Moderate
$1 < \mathrm{BF}_{10} \leq 3$	Anecdotal
For the null model	
$\frac{1}{3} < \mathrm{BF}_{10} \leq 1$	Anecdotal
$\frac{1}{10} < \mathrm{BF}_{10} \leq \frac{1}{3}$	Moderate

Table 2: Criteria for interpreting Bayes factors (Lee and Wagenmakers, 2013, derived from Jeffreys, 1939/1998, excerpt relevant to the current study)

Given the substantial susceptibility of Bayes factors to prior settings for the explanatory variables and intercept (Nicenboim et al., to appear), we conducted prior predictive checks to calibrate the priors for intercepts, explanatory variables, and covariates, following Schad et al's (2020a; 2022) methodologies. Moreover, we calculated BF_{10} iteratively for each explanatory variable using normally-distributed priors with a mean of zero and a range of standard deviations (Nicenboim et al., 2020; Ogawa, 2023). This approach allowed us to observe the trends in BF_{10} and coefficients across different prior specifications. See Appendix C. for further details.

3.5.1 Reading time

We modelled the reading times using a log-normal distribution. The key explanatory variables were:

- the target voice (V-∅ active or V-(*r*)*are* passive)
- the verb class difference for each target voice
  - =*ni*$_{\text{DAT}}$-verbs or =*o*$_{\text{ACC}}$-verbs in active voice
  - =*ni*$_{\text{DAT}}$-verbs or =*o*$_{\text{ACC}}$-verbs in passive voice.

Sum-coding was applied to the target voice variable, and nested sum-coding to the verb class differences (Schad et al., 2020b). The covariates in the model included the number of characters in the region and the absolute trial order, both of which were standardised (Nicenboim et al., to appear). Details are provided in Appendix C.

3.5.2 Comprehension accuracy

Accuracy of the comprehension questions was analysed with mixed effects logistic regressions. We focused on seven key explanatory variables:

- the target voice
- priming (match versus mismatch in voice between target and comprehension question)
- the interaction of the two factors above
- the verb class difference for each target voice and priming
  - active =*ni*$_{\text{DAT}}$-verbs versus =*o*$_{\text{ACC}}$-verbs in both target and question
  - passive =*ni*$_{\text{DAT}}$-verbs versus =*o*$_{\text{ACC}}$-verbs in both target and question
  - =*ni*$_{\text{DAT}}$-verbs versus =*o*$_{\text{ACC}}$-verbs in active target and passive question
  - =*ni*$_{\text{DAT}}$-verbs versus =*o*$_{\text{ACC}}$-verbs in passive target and active question

The first three variables were sum-coded and the rest were nested sum-coded. The z-transformed absolute trial order was also included as a covariate. Further details can be found in in Table 5 in Appendix C.

3.6 Predictions

3.6.1 Reading time

If a =*ni*$_{\text{DAT}}$-marked NP strongly predicts passives in Japanese and such predictions facilitate the reading of passives, shorter reading times for passives could be observed in the verb region (R5). Furthermore, if the parser begins constructing the passive structure in R3 or immediately after in R4 due to the presence of a $=ni_{\text{DAT}}$-marked NP, longer reading times may also occur in these regions.

However, only in the active $=ni_{\text{DAT}}$-verb condition, the presence of a $=ni_{\text{DAT}}$-marked NP would mislead the parser into anticipating a passive sentence. This would cause surprisal and longer reading times in R5 of $=ni_{\text{DAT}}$-verbs, as the actual sentence turns out to be active in that region.

It is, nonetheless, also unsurprising to find longer reading times in passives in both $=ni_{\text{DAT}}$- and $=o_{\text{ACC}}$-verbs, as even when the morphological structure of verbs is matched as closely as possible, passive verbs in Japanese may still result in longer reading times (Ogawa, 2023).

We may also observe the same reading time pattern at R6, due to a spill-over and/or delay of the processing cost from the verb region (R5).

3.6.2 Comprehension accuracy

As priming effects were found both between active targets and questions, and between passive targets and questions (Ogawa, 2023), higher accuracy is expected when target and the question share the same voice, and lower accuracy when they not.

If, in addition, a $=ni_{\text{DAT}}$-marked NP serves as a predictor for passive sentences, the prediction of a passive structure could facilitate more accurate comprehension of passive targets. Thus, even in the passive condition, accuracy is expected to be as high as in the active condition. However, in the active $=ni_{\text{DAT}}$-verb condition, the parser may initially predict a passive structure at R3 but then realize at R5 that the sentence is actually active. This could lead to surprisal, resulting in a significant drop in accuracy specifically in this condition.

4 Results

4.1 Longer reading times for passives

V-(*r*)*are* passives elicited longer median and mean reading times than V-∅ actives, especially in the verb region (R5), as shown in Table 3. As highlighted in Figure 5 in Appendix D., Bayes factor analyses indicate moderate to very strong evidence in support of the effect of voice. These results align with the previous finding of increased reading times for Japanese passives (Ogawa, 2023).

However, no significant differences in reading times were found between actives and passives in R3 and R4. Bayes factors for these regions were below 1, signifying an absence of the voice effect. Therefore, it remains inconclusive whether the parser actively predicts passive constructions upon reading the case marker $=ni_{\text{DAT}}$.

Interestingly, when comparing reading times of R6 between active $=ni_{\text{DAT}}$-verb condition and active $=o_{\text{ACC}}$-verb condition, the reading times were longer after active $=ni_{\text{DAT}}$-verbs, and Bayes factors indicate moderate evidence supporting a difference. This suggests that in the active $=ni_{\text{DAT}}$-verb condition, the parser may initially predict a passive structure at R3 by $=ni_{\text{DAT}}$ but recognise at R5 that the sentence is indeed active, leading to a delayed reanalysis at R6.

Voice	Verb class	R3: NP2 Median (Mean)	R4: ADV Median (Mean)	R5: Verb Median (Mean)	R6: Modal Median (Mean)
V-∅ active	=*ni*-verb	800 (1143.9)	664 (940.4)	823 (1088.5)	526 (696.4)
	=*o*-verb	754.5 (1069.9)	647.5 (858)	916 (1141.8)	508.5 (619.7)
V-(*r*)*are* passive	=*ni*-verb	816 (1142)	648.5 (877.1)	1120.5 (1528.7)	543 (753.8)
	=*o*-verb	752 (1101.5)	679 (977.5)	1093 (1544.4)	538 (716.3)

Table 3: Median and mean reading time (ms) by condition

4.2 Lower comprehension accuracy for passives

Figure 1 illustrates that, overall, accuracy is lower for passives compared to actives. It also shows that accuracy is higher when the voice of the target sentence matches that of the corresponding question, regardless of whether the target is active or passive. This result is strongly supported by Bayes factors, which provide moderate to extreme evidence, as shown in Figure 2.

Paolazzi et al. (2021b) discussed the increased accuracy in passive sentence comprehension when both the target sentence and the question are passive. Our results support this finding and also demonstrate that comprehension accuracy is higher when both the target and the question are active. This phenomenon, where accuracy is higher when the voice of the target sentence matches that of the question, is independent of the voice, corroborating earlier findings (Ogawa, 2023).

However, the differences between $=o_{\text{ACC}}$-verbs and $=ni_{\text{DAT}}$-verbs, regardless of voice or priming conditions, were not supported by Bayes factor analysis. In fact, the Bayes factors consistently fell below 1. Therefore, there is no significant benefit to passive sentence comprehension from the case marker itself.

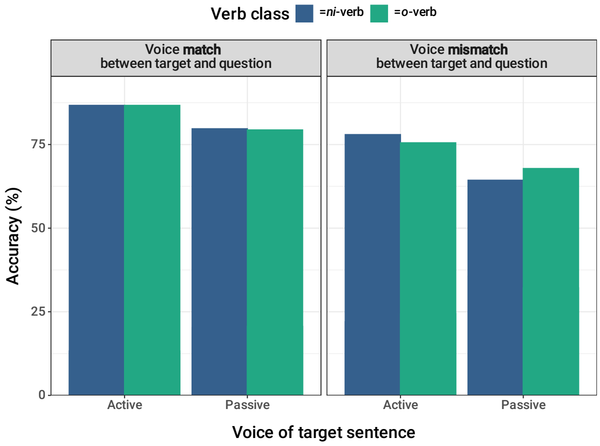


Figure 1: Raw accuracy for the comprehension question by condition

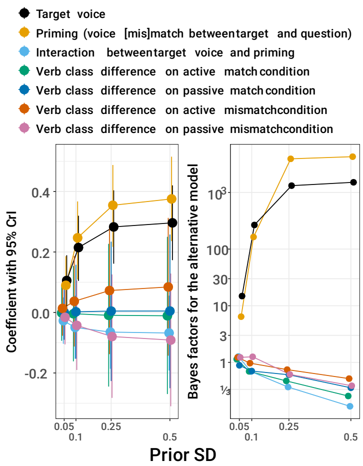


Figure 2: Change in estimates (with 95% Credible Interval) and Bayes factor for factors by prior SD

5 General discussion and conclusion

5.1 Predicting a passive construction from a case marker may be difficult

Our main purpose in this study was to determine whether the human parser can predict passive constructions from linguistic elements preceding the sentence-final verb in Japanese, before confirming this by reading the verb. We hypothesized that case markers such as =*o*$_{\text{ACC}}$ and =*ni*$_{\text{DAT}}$ function as passive predictors. To test this, we conducted an SPR experiment tracking reading times and examined accuracy through comprehension questions.

The results did not provide evidence that differences in case markers lead to faster reading times or higher accuracy for passive sentences. However, similar to previous research on Japanese passives (Ogawa, 2023), we demonstrated increased reading times and lower accuracy for verbs in passives. Unlike Ogawa (2023), we found strong evidence through Bayes factor analysis for this increase in reading time. It is important to note that while Ogawa (2023) controlled for morphological complexity by using V-*te morau* benefactive passive and V-*te ageru* benefactive active, our study used pairs of V-(*r*)*are* passive and V-∅ active, which differ in morphological structure and character count. This discrepancy may have contributed to the statistically significant results. Yet, given that character count was a covariate in our statistical models, the increased reading times for passive sentences cannot be solely explained by morphological complexity or word length.

Based on previous research, which suggests that case markers preceding verbs can help the human parser predict sentence structures (Muraoka, 2006), the current reading time results could be interpreted as indicating that the case marker =*ni*$_{\text{DAT}}$ contributes to predicting ditransitive constructions rather than passives. This is because =*ni*$_{\text{DAT}}$ is used in ditransitive constructions (e.g., NP=*ga*$_{\text{NOM}}$ NP=*ni*$_{\text{DAT}}$ NP=*o*$_{\text{ACC}}$ V), as well as passives. Therefore, the parser might find it difficult to predict passive sentences solely from the presence of =*ni*$_{\text{DAT}}$.

In fact, Muraoka (2006)'s results (see Figure 3 in Appendix A.) show that an accusative NP forming ditransitive sentences (211 occurrences) is predicted more frequently than a passive verb (45 occurrences) immediately following =*ni*$_{\text{DAT}}$. Consequently, if the human parser predicts that the sentence is a ditransitive construction upon encountering =*ni*$_{\text{DAT}}$ and expects an =*o*$_{\text{ACC}}$-NP to follow, the presence of a passive verb (R5) would cause surprisal, as it indicates that the expected ditransitive construction is impossible. However, in the active $=ni_{\text{DAT}}$-verb condition, the reading times in R5 were as short as those in the active $=o_{\text{ACC}}$-verb condition, even that condition was also against the prediction of ditransitives. Therefore, it cannot be conclusively stated that $=ni_{\text{DAT}}$ primarily predicts ditransitive constructions. Rather, it is also possible that the case marker $=ni_{\text{DAT}}$ does not efficiently contribute to predicting either passive or ditransitives.

It is worth considering that the sentence completion task in Muraoka (2006) (which involves both comprehension and production) and the current SPR experiment (which is comprehension-oriented) might differ in their sensitivity to detecting the prediction of elements that follow $=ni_{\text{DAT}}$. Future SPR experiments comparing reading times of the regions following $=ni_{\text{DAT}}$ in ditransitive and passive sentences could provide a more precise understanding of the case marker's role in the prediction during sentence comprehension.

5.2 Priming influences comprehension accuracy for both passives and actives

Regarding accuracy for passive sentences, we observed a robust priming effect: accuracy was higher when the voice of the target sentence matched that of the corresponding question, regardless of whether the target was active or passive. Despite this priming effect, overall comprehension of passive sentences remained lower compared to active sentences.

As shown in Figure 2, Bayes factor analyses indicated that there was no difference between the effect of voice match versus voice mismatch within actives and the effect of voice match versus voice mismatch within passive sentences (i.e., no significant interaction between target voice and priming). This suggests that both actives and passives are equally error-prone when the voice of the target sentence and the comprehension question differ.

Previous structural priming research using SPR experiments has shown that a less frequent construction is more primable (Wei et al., 2016). Given that passives are less frequent than actives in Japanese (Aoyama, 2023), passives would be more primable, leading to a larger difference in accuracy between voice-matched and voice-mismatched conditions for passives compared to actives. However, Ogawa (2023)'s experiment comparing benefactive passives and benefactive actives found that both constructions were error-prone when there was a voice mismatch, and our experiment replicated this finding. Therefore, these studies suggest that, contrary to previous research, the priming effect may be more robust than construction frequency.

Acknowledgements

I extend my sincere gratitude to Kei Furukawa, Chuyu Huang, Takeshi Kishiyama, and Itsuki Minemi for their insightful discussion. Their feedback significantly enhanced the clarity of the current paper. I am also heavily indebted to Yue Teng for improving the usability of my PCIbex application. The author affirms that this acknowledgment is expressed with the prior consent of all the aforementioned colleagues.

Research funding for this study was provided by Grant-in-Aid for JSPS Fellows Grant Numbers JP19J21705.

Lastly, I wish to express my appreciation to all participants who engaged in the experiment.

References

Gerry T.M. Altmann. 2011. Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta Psychologica*, 137(2):190–200.

Tatsuya Aoyama. 2023. Corpus-based investigation of the markedness and frequency of Japanese passives in contemporary written Japanese. *Society for Computation in Linguistics*, 6(1):361–363.

Edward Keith Brown and Anne H. Anderson. 2006. Encyclopedia of language & linguistics.

Paul-Christian Burkner. 2021. *brms: Bayesian Regression Models using Stan*. R package version 2.16.3.

Department of Linguistics of Max Planck Institute for Evolutionary Anthropology. 2008. The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses.

Jonah Gabry and Rok Češnovar. 2021. *cmdstanr: R Interface to CmdStan*. Https://mc-stan.org/cmdstanr.

Nino Grillo, Artemis Alexiadou, Berit Gehrke, Nils Hirsch, Caterina Paolazzi, and Andrea Santi. 2019. Processing unambiguous verbal passives in german. *Journal of Linguistics*, 55(3):523–562.

Quentin F. Gronau and Henrik Singmann. 2021. *bridgesampling: Bridge Sampling for Marginal Likelihoods and Bayes Factors*. R package version 1.1-2.

Jiqiang Guo, Jonah Gabry, Ben Goodrich, and Sebastian Weber. 2021. *rstan: R Interface to Stan*. R package version 2.21.3.

Harold Jeffreys. 1939/1998. *The Theory of Probability*. Oxford University Press.

Marcel A. Just, Patricia A. Carpenter, and Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2):228–238.

Taro Kageyama. 1993. *Bunpō to Gokēsei [Grammar and word formation]*. Hituzi Syobo, Tokyo.

Taro Kageyama. 2013. Word structure. In *Compound Verb Lexicon*. National Institute for Japanese Language and Linguistics (NINJAL).

Ryuta Kinno, Mitsuru Kawamura, Seiji Shioda, and Kuniyoshi L. Sakai. 2008. Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Human Brain Mapping*, 29(9):1015–1027.

Masatoshi Koizumi and Satoshi Imamura. 2017. Interaction between syntactic structure and information structure in the processing of a head-final language. *Journal of Psycholinguistic Research*, 46(1):247–260.

Michael David Lee and Eric-Jan Wagenmakers. 2013. *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Michael Meng and Markus Bader. 2020. Does comprehension (sometimes) go wrong for noncanonical sentences? *Quarterly Journal of Experimental Psychology*, 74(1):1–28.

Satoru Muraoka. 2006. The effects of case marking information on processing object nps in japanese. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 13(3):404–416.

National Institute for Japanese Language and Linguistics and Lago Institute of Language. 2012. *NINJAL-LWP for BCCWJ*.

Bruno Nicenboim, Daniel Schad, and Shravan Vasishth. to appear. *An Introduction to Bayesian Data Analysis for Cognitive Science*. CRC Press.

Bruno Nicenboim, Shravan Vasishth, and Frank Rösler. 2020. Are words pre-activated probabilistically during sentence comprehension? evidence from new data and a bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, 142:107427.

Masataka Ogawa. 2023. Japanese benefactive passives are difficult to comprehend than benefactive actives. In *Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation*, pages 174–193, Hong Kong, China. Association for Computational Linguistics.

Dario Paape, Shravan Vasishth, and Ralf Engbert. 2021. Does local coherence lead to targeted regressions and illusions of grammaticality? *Open Mind*, page 1–17.

Caterina Laura Paolazzi, Nino Grillo, Artemis Alexiadou, and Andrea Santi. 2016. Processing English passives: Interaction with event structure, but no evidence for heuristics. In *29th Annual CUNY Conference on Human Sentence Processing*, University of Florida.

Caterina Laura Paolazzi, Nino Grillo, Artemis Alexiadou, and Andrea Santi. 2019. Passives are not hard to interpret but hard to remember: evidence from online and offline studies. *Language, Cognition and Neuroscience*, 34(8):991–1015.

Caterina Laura Paolazzi, Nino Grillo, Claudia Cera, Fani Karageorgou, Emily Bullman, Wing Yee Chow, and Andrea Santi. 2021a. Eyetracking while reading passives: an event structure account of difficulty. *Language, Cognition and Neuroscience*, 37(2):135–153.

Caterina Laura Paolazzi, Nino Grillo, and Andrea Santi. 2017. Passives are not always more difficult than actives. In *Proceedings of the Architectures and Mechanisms for Language Processing 2017*. AMLaP.

Caterina Laura Paolazzi, Nino Grillo, and Andrea Santi. 2021b. The source of passive sentence difficulty: Task effects and predicate semantics, not argument order. In *Passives Cross-Linguistically*, pages 359–393. Brill.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Daniel J. Schad, Michael Betancourt, and Shravan Vasishth. 2020a. Toward a principled Bayesian workflow in cognitive science. *Psychological Methods*.

Daniel J. Schad, Bruno Nicenboim, Paul-Christian Bürkner, Michael Betancourt, and Shravan Vasishth. 2022. Workflow techniques for the robust use of Bayes factors. *Psychological Methods*.

Daniel J. Schad, Shravan Vasishth, Sven Hohenstein, and Reinhold Kliegl. 2020b. How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110:104038.

Katsuo Tamaoka, Hiromu Sakai, Jun-ichiro Kawahara, Yayoi Miyaoka, Hyunjung Lim, and Masatoshi Koizumi. 2005. Priority information used for the processing of japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research*, 34(3):281–332.

Kyohei Tanaka, Shinri Ohta, Ryuta Kinno, and Kuniyoshi L. Sakai. 2017. Activation changes of the left inferior frontal gyrus for the factors of construction and scrambling in a sentence. *Proceedings of the Japan Academy, Series B*, 93(7):511–522.

Tasaku Tsunoda. 1985. Remarks on transitivity. *Journal of Linguistics*, 21(2):385–396.

Tasaku Tsunoda. 2009. *Sekai no Gengo to Nihongo: Gengo-ruikeiron kara mita Nihongo [Languages of the world and Japanese language: Japanese languade from typological perspectives]*, pages 47–53. Kurosio Publishers, Tokyo. Japanese.

Hang Wei, Yanping Dong, Julie E. Boland, and Fang Yuan. 2016. Structural priming and frequency effects interact in Chinese sentence comprehension. *Frontiers in Psychology*, 7.

Jeffrey D. Witzel and Naoko O. Witzel. 2011. The processing of Japanese control sentences. In Hiroko Yamashita, Yuki Hirose, and Jerome L. Packard, editors, *Processing and Producing Head-final Structures*, pages 23–47. Springer Netherlands, Dordrecht.

Satoru Yokoyama, Tadao Miyamoto, Jorge Riera, Jungho Kim, Yuko Akitsuki, Kazuki Iwata, Kei Yoshimoto, Kaoru Horie, Shigeru Sato, and Ryuta Kawashima. 2006. Cortical Mechanisms Involved in the Processing of Verbs: An fMRI Study. *Journal of Cognitive Neuroscience*, 18(8):1304–1313.

Appendix A. Elements filled in the sentence completion task by Muraoka (2006)

Figure 3 indicates the token frequency of elements that participant filled in the sentence completion task by Muraoka (2006). Participants produced passivised verbs (45 occurrences) after they saw $=_{\text{DAT}}$-marked NPs, whereas they produced 211 accusative NPs to form ditransitive sentences.

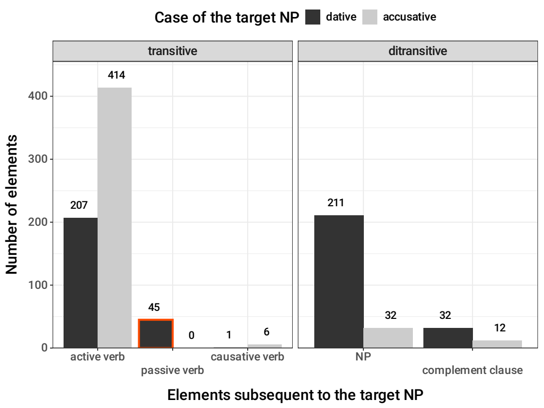


Figure 3: Elements filled in the sentence completion task by Muraoka (2006, pp.407–408, Experiment 1). Visualisation mine.

Appendix B. Sample of comprehension questions

Sample for the comprehension question (*naguritobas-u* 'punch')

c. **Active question in NP1 → NP2 order**
('Yes' for V-∅ / 'No' for V-(*r*)*are* / 'match' to V-∅ in priming / 'mismatch' to V-(*r*)*are* in priming)

Takahashi=ga Ōtsuka=o naguritobashi-ta-rashī-desu-ka?
T.=NOM Ō.=ACC hit-PST-INFR-POL-Q
Does Takahashi seem to have punched Ōtsuka?

d. **Active question in NP2 → NP1 order**
('No' for V-∅ / 'Yes' for V-(*r*)*are* / 'match' to V-∅ in priming / 'mismatch' to V-(*r*)*are* in priming)

Ōtsuka=ga Takahashi=o naguritobashi-ta-rashī-desu-ka?
Ō.=NOM T.=ACC hit-PST-INFR-POL-Q
Does Ōtsuka seem to have punched Takahashi?

e. **Passive question in NP1 → NP2 order**
('No' for V-∅ / 'Yes' for V-(*r*)*are* / 'mismatch' to V-∅ in priming / 'match' to V-(*r*)*are* in priming)

Ōtsuka=ga Takahashi=ni naguritobas-are-ta-rashī-desu-ka?
Ō.=NOM T.=DAT hit-PASS-PST-INFR-POL-Q
Did Ōtsuka seem to have been punched by Takahashi?

f. **Passive question in NP2 → NP1 order**
('Yes' for V-∅ / 'No' for V-(*r*)*are* / 'mismatch' to V-∅ in priming / 'match' to V-(*r*)*are* in priming)

Takahashi=ga Ōtsuka=ni naguritobas-are-ta-rashī-desu-ka?
T.=NOM Ō.=DAT hit-PASS-PST-INFR-POL-Q
Did Takahashi seem to have been punched by Ōtsuka?

Appendix C. Contrasts to code explanatory variables and priors used in the current study

Reading time

Our key explanatory variables for reading time data are the following three factors: the target voice (V-

∅ active versus V-(*r*)*are* passive) and the verb class difference for each target voice (=ni_{DAT}-verbs versus =o_{ACC}-verbs in active voice [active.o.vs.ni], and =ni_{DAT}-verbs versus =o_{ACC}-verbs in passive voice [passive.o.vs.ni]). We sum-coded the target voice, and for each target voice, we coded the verb class difference using nested sum contrast: =ni_{DAT}-verbs versus =o_{ACC}-verbs in active voice (active.o.vs.ni), and =ni_{DAT}-verbs versus =o_{ACC}-verbs in passive voice (passive.o.vs.ni), as shown below.

$$\text{voice} = \begin{cases} 1 & \text{(passive)} \\ -1 & \text{(active)} \end{cases}$$

$$\text{active.o.vs.ni} = \begin{cases} 1 & \text{(active } =ni_{\text{DAT}}\text{-verb)} \\ 0 & \text{(passive verbs)} \\ -1 & \text{(active } =o_{\text{ACC}}\text{-verb)} \end{cases}$$

$$\text{passive.o.vs.ni} = \begin{cases} 1 & \text{(passive } =ni_{\text{DAT}}\text{-verb)} \\ 0 & \text{(active verbs)} \\ -1 & \text{(passive } =o_{\text{ACC}}\text{-verb)} \end{cases}$$

According to prior predictive checks, we used the following priors for target voice and the verb class difference for each target voice: $\mathrm{N}(0, 0.5)$, $\mathrm{N}(0, 0.25)$, $\mathrm{N}(0, 0.1)$, $\mathrm{N}(0, 0.075)$, $\mathrm{N}(0, 0.05)$, $\mathrm{N}(0, 0.025)$, $\mathrm{N}(0, 0.01)$, $\mathrm{N}(0, 0.0075)$, $\mathrm{N}(0, 0.005)$, $\mathrm{N}(0, 0.0025)$, $\mathrm{N}(0, 0.001)$. Table 4 shows priors for other parameters.

Coefficient	R3: Second NP	R4: ADV on action	R5: Verb	R6: Modal particle
Intercept	$\mathrm{N}(6.7, 0.1)$	$\mathrm{N}(6.5, 0.2)$	$\mathrm{N}(6.9, 0.2)$	$\mathrm{N}(6.3, 0.1)$
Region length	(Not used in the model)	$\mathrm{N}(0, 0.1)$	$\mathrm{N}(0, 0.1)$	$\mathrm{N}(0, 0.1)$
Trial order	$\mathrm{N}(0, 0.05)$	$\mathrm{N}(0, 0.05)$	$\mathrm{N}(0, 0.1)$	$\mathrm{N}(0, 0.01)$
Scale parameter σ	$\mathrm{N}_+(0, 0.2)$	$\mathrm{N}_+(0, 0.4)$	$\mathrm{N}_+(0, 0.1)$	$\mathrm{N}_+(0, 0.2)$
Parameters for random effects				
SD τ	$\mathrm{N}(0, 0.2)$	$\mathrm{N}(0, 0.1)$	$\mathrm{N}(0, 0.1)$	$\mathrm{N}(0, 0.2)$
Correlation parameter ρ	$\mathrm{LKJ}(\eta = 2)$	$\mathrm{LKJ}(\eta = 2)$	$\mathrm{LKJ}(\eta = 2)$	$\mathrm{LKJ}(\eta = 2)$

Table 4: Priors used to analyse reading time data

Comprehension accuracy

Table 5 illustrates how we coded our seven key explanatory variables.

Based on prior predictive checks, we used the following priors for target voice and the verb class difference for each target voice: $\mathrm{N}(0, 0.5)$, $\mathrm{N}(0, 0.25)$, $\mathrm{N}(0, 0.1)$, $\mathrm{N}(0, 0.05)$. We used $\mathrm{N}(1.3, 0.2)$ priors for intercepts, $\mathrm{N}(0, 0.1)$ priors for the slopes, and LKJ priors with $\eta = 2$ for the correlation matrices.

Appendix D. Raw reading times in the self-paced reading (SPR) task

Figure 4 shows the raw reading times for each region in our SPR task by condition.

Appendix E. Coefficient and Bayes factors for each key explanatory variables on reading time difference

Figure 5 to Figure 7 illustrate the estimated coefficient and Bayes factors for each key explanatory variables on reading time difference in our SPR task.

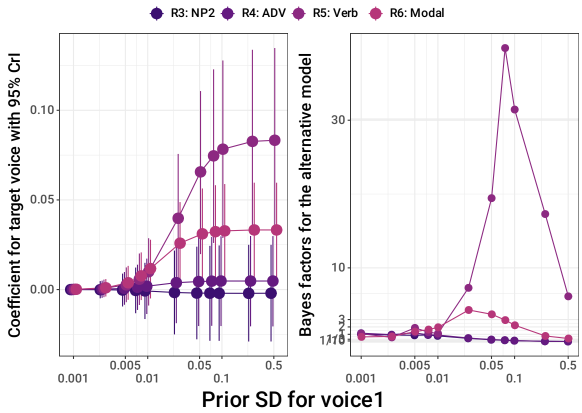


Figure 5: Change in the estimate of voice effect (with 95% Credible Interval) and Bayes factor for voice by prior SD in the regions of NP2 (R3), ADV (R4), verb (R5), and the modal (R6)

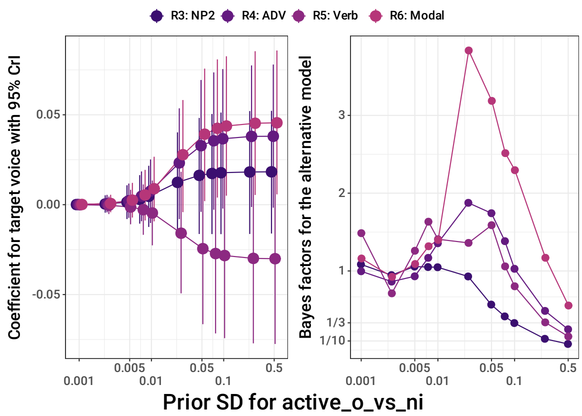
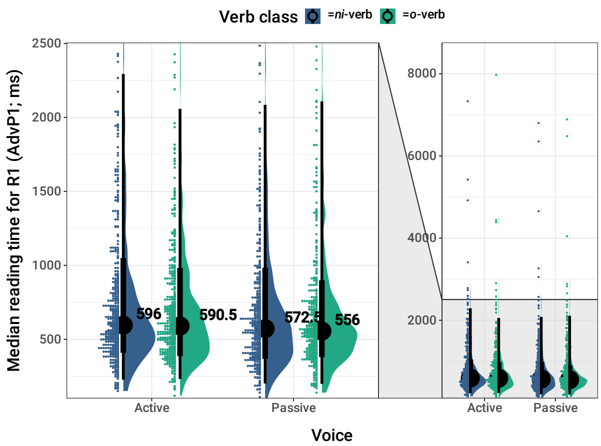
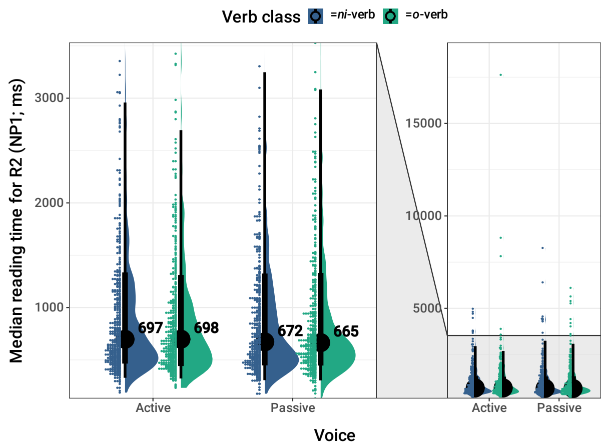


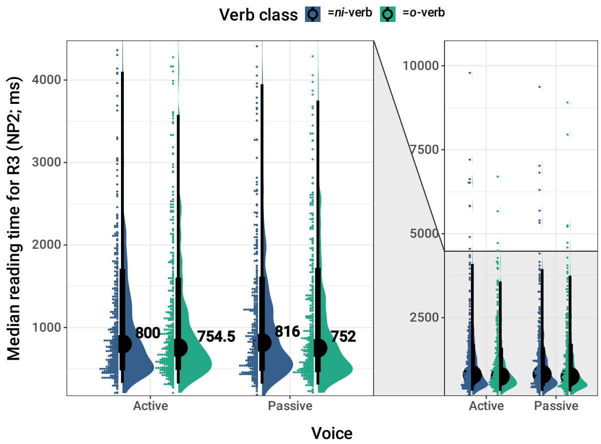
Figure 6: Change in the estimate of difference between active =*o*-verb and active =*ni*-verb (with 95% Credible Interval) and Bayes factor for verb class difference in active by prior SD in the regions of NP2 (R3), ADV (R4), verb (R5), and the modal (R6)



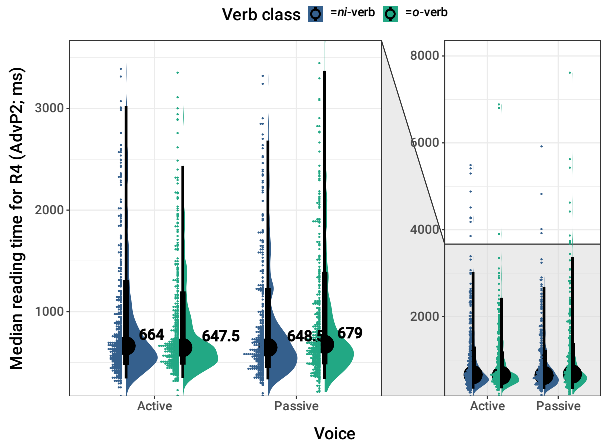
(a) R1: Adverb phrase



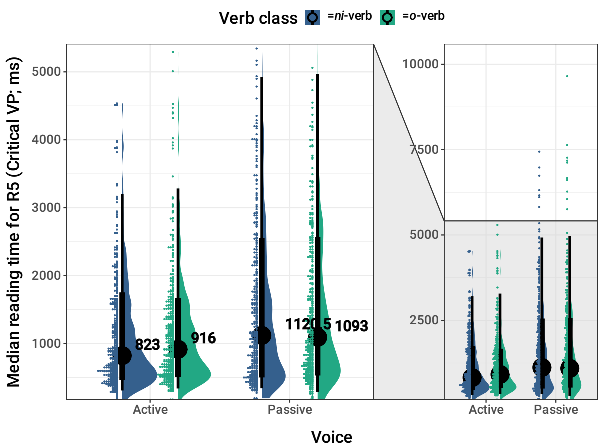
(b) R2: First NP



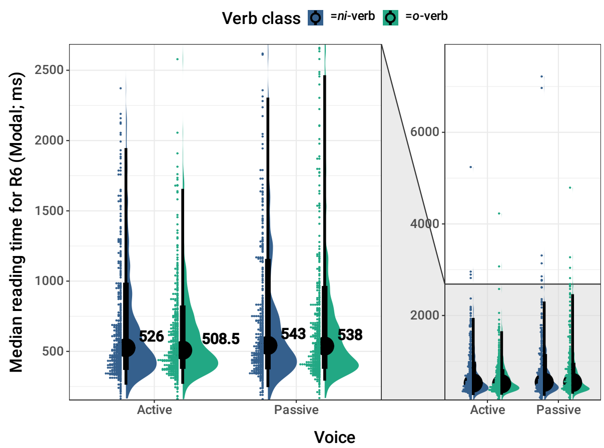
(c) R3: Second NP



(d) R4: Adverb on action



(e) R5: Verb region



(f) R6: Modal particle region

Figure 4: Raw reading time for each region; Thick bars and thin bars indicate the 66% and 95% quantile intervals of data respectively, and bullets indicate the median reading time.

condition			Explanatory variables in the models						
Voice	Priming	Case pattern	voice	priming	voice: priming	active. match. o.vs.ni	active. mismatch. o.vs.ni	passive. match. o.vs.ni	passive. mismatch. o.vs.ni
passive	match	*=o*	1	-1	-1	0	0	-1	0
		=ni	1	-1	-1	0	0	1	0
	mismatch	*=o*	1	1	1	0	0	0	-1
		=ni	1	1	1	0	0	0	1
active	match	*=o*	-1	-1	1	-1	0	0	0
		=ni	-1	-1	1	1	0	0	0
	mismatch	*=o*	-1	1	-1	0	-1	0	0
		=ni	-1	1	-1	0	1	0	0

Table 5: Coding for the explanatory variables for reaction time of the correctly answered comprehension questions in Experiment

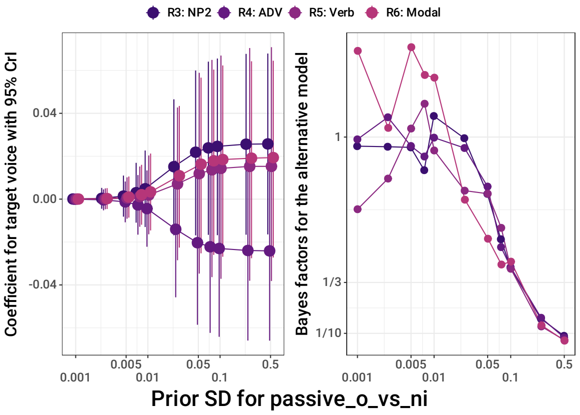


Figure 7: Change in the estimate of difference between passive =*o*-verb and passive =*ni*-verb (with 95% Credible Interval) and Bayes factor for verb class difference in passive by prior SD in the regions of NP2 (R3), ADV (R4), verb (R5), and the modal (R6)

<!–