No Revision Required, Still Difficult to Interpret:  
Japanese Children’s Comprehension of Verb-Initial Passives  

Megumi Ishikawa, Takane Ito, and Takuya Goro

1. Introduction

In first language acquisition research, it has been repeatedly observed that 4- and 5-year-old children showed poor performance with tasks involving transitive full passive comprehension. Children from various different linguistic communities showed the same difficulties, including Japanese speaking children (e.g., Armon-Lotem et al., 2016; Harris, 1976; Maratsos et al., 1985; Okabe & Sano, 2002, among others). For example, while Japanese 4-year-olds have no problem in interpreting transitive active sentences such as (1a), their performance with the full passive counterpart (1b) is somewhat degraded, which holds across different interpretive tasks (e.g., Truth Value Judgment, Picture-Selection, etc.)

(1) a)  
Japanese Active  
Onnanoko-ga otokonoko-o ker-ta-yo  
Girl-NOM boy-ACC kick-PAST-PART  
‘The girl kicked the boy.’

b)  
Japanese Full Passive  
Otokonoko-ga onnanoko-ni ker-are-ta-yo  
Boy-NOM girl-by kick-PASS-PAST-PART  
‘The boy was kicked by the girl.’

* Megumi Ishikawa (ishikawa.mg@gmail.com) and Takane Ito, The university of Tokyo, Takuya Goro, Tsuda University. We would like to thank all the participants of our experiment as well as the teachers of Takachiho kindergarten, and five anonymous BUCLD reviewers and the audience at BUCLD for valuable comments. This research was in part supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Kakenhi) to the 1st author (#15J09804) and to the 2nd author (#16H03429). I sincerely thank the Diversity Enhancement Award for BUCLD 44th. Any errors and shortcomings are solely our own.

1 List of abbreviations: NOM = nominative case marker, ACC = accusative case marker, DAT = dative case marker, PASS = passive morpheme, PAST = past tense, and PART = particle.
One major approach to children’s difficulty with full passive comprehension is grammar-based accounts (Babyonishev, 2001; Borer & Wexler, 1987; Fox & Grodzinsky, 1998; Hyams & Snyder, 2005, 2006; Orfitelli, 2012; Wexler, 2004, among others). These accounts differ from each other with respect to theoretical details; therefore, they have various different empirical consequences, which we will not examine any further in this paper. For our present purpose, it suffices to note that they all share the assumption that something in children’s grammar is the source of the difficulty with full passive comprehension. Instead, we would like to focus on a recently emerging alternative to the grammar-based approach: the Incremental Processing Hypothesis (IPH: Deen et al., 2018; Huang et al., 2013). In a nutshell, the IPH is an attempt to reduce children’s difficulty with passives to independently motivated properties of children’s sentence processing. According to the IPH, the source of children’s difficulty in full passive comprehension lies in their problem in revising the initial parse of incoming sentences.

The current study is an attempt to assess empirical predictions of the IPH, to test if it can really replace the grammatical accounts. We report the results of an experiment with Japanese-speaking 4- and 5-year-old children. The experiment utilized the flexible word-order in Japanese to create a novel set of test conditions. We begin by reviewing the IPH account of children’s difficulty with passives.

2. The IPH account of the passive difficulty

The IPH, inspired by theories on incremental sentence processing in adults (Trueswell & Tanenhaus, 1994), consists of the following three key ideas. Each of the three key ideas is independently motivated by observations made in various studies of sentence processing of adults and children (Bever, 1970; Omaki et al., 2014; Tanenhaus et al., 1995; Trueswell et al., 1999; among others).

(2) The key ideas of the Incremental Processing Hypothesis (IPH)
   a) Agent first bias: children have a bias to assign an agent role to the first animate NP in a sentence.
   b) Incremental sentence processing: children process sentences incrementally.
   c) Revision failure: children often fail to revise their initial parse of a sentence.

First, young children have a bias to map the first (animate) NP in a sentence to an agent role, probably due to the greater probability of active sentences in the input (Bever, 1970). Second, this agent-role assignment is reported to happen incrementally: children do not wait for the main predicate to determine the thematic role of the NPs that appear before the predicate (e.g., Abbot-Smith et al., 2017). Evidence for children’s incremental sentence processing originally came from their non-adult garden-path interpretations. For example, in the study by Trueswell et al. (1999), 4- and 5-year-old English-speaking children heard the test
sentences that involve local PP-attachment ambiguity, such as *Put the frog on the napkin in the box*. When asked to act-out the test sentence, many children showed a considerable confusion, and their non-adult responses suggested that they interpreted the PP *on the napkin* as the destination of the frog. This observation is interpreted as showing that children incrementally interpreted the PP *on the napkin* as the goal argument of the verb *put* (=2b), and that they failed to revise this initial parse even after encountering the disambiguating PP *in the box* (=2c). Thus, children are more prone to be garden-pathed than adults: they often fail to revise the initial incremental parse, which leads them to adopt apparent misinterpretations in experimental settings.

Let us now consider how the IPH (2) can be applied to Japanese passives to explain children’s difficulty with them. The canonical word order in Japanese is SOV, and therefore the passive morpheme -(r)are appears in the sentence-final position.

(3) Japanese Full Passive

Tora-ga zou-ni ker-are-ta-yo
Tiger-NOM elephant-by kick-PASS-PAST-PART
‘The tiger was kicked by the elephant.’

The -ni marked NP corresponds to the by-phrase in English passives. However, the particle -ni is ambiguous, and is also used as a locative P marker / a dative case in active sentences:

(4) Locative -ni

Zou-ga kouen-ni ik-ta-yo
Elephant-NOM park-to go-PAST-PART
‘The elephant went to the park.’

(5) Dative -ni

Zou-ga tora-ni tegami-o watasi-ta-yo
Elephant-NOM tiger-DAT letter-ACC pass-PAST-PART
‘The elephant passed the tiger a letter.’

This ambiguity of -ni entails that the presence of the ni-marked NP does not allow the parser to determine whether the sentence is an active or a passive. In fact, Hirotani et al. (2011) reported that Japanese adults incrementally processed a “NP-ga NP-ni” sequence as a part of an active sentence.

Following the IPH, let us assume that Japanese children hearing the sentence in (3) initially map the first animate NP *Tora-ga ‘Tiger’* onto an agent role. The second NP, *zou-ni*, is then probably mapped onto a goal role, assuming that children have an adult-like tendency in interpreting a NP-ni. The parser has to wait until the sentence-final position for the disambiguating cue: the passive morpheme -(r)are. The presence of -(r)are requires revision of the initial incremental theta-assignments. However, if children’s parser fails to execute the
revision, then they would essentially be garden-pathed, which would lead to non-adult pattern of responses in experimental settings.

Under the IPH account, children’s difficulties in comprehending full passive sentences stem from their difficulties in revising the initial parsing commitments. Since children’s revision failures have been observed with a variety of constructions and experimental conditions (Choi & Trueswell, 2010; Hurewitz et al., 2000; Weighall, 2008; Omaki et al., 2014), it is reasonable to assume that the tendency to stick to the initial incremental commitments is a general property of child sentence processing. The difficulty with passives, then, can now be regarded as a consequence of such general properties.

This makes the IPH a conceptually attractive alternative to the traditional grammatical accounts. The IPH can account for the children’s problem with full passive comprehension without any stipulations on child grammar specific to the passive construction. Since there are independent empirical grounds for assuming revision failure in child parser, it would be desirable if the passive difficulty is reduced to the general property. Given this conceptual advantage of the IPH, it now becomes important to examine empirical consequences of the hypothesis.

Several previous studies argued for the IPH, based on their experimental evidence. Deen et al. (2018) found that English-speaking children performed better on a Truth-Value Judgment Task when the passive test sentence was repeated (i.e., presented twice) before the child made her judgment. Huang et al. (2013) observed that Mandarin-speaking children were more successful if the passive marker occurred before the first possible theta-role assignment. These findings suggest that children’s degraded comprehension of passives is (at least partly) due to processing factors. In an attempt to supplement this line of empirical research, we carried out an experiment with Japanese-speaking children. We sought to find out how children fare with passive sentences when the passive marker occurs before any NPs in the sentence. The flexible word-order in Japanese allowed us to construct the crucial sentences to investigate the question: the V-initial sentences.

3. Japanese V-initial Sentences

As briefly mentioned in the previous section, the canonical word order in Japanese is SOV. However, it is also possible to bring the verb to the sentence-initial position. This word-order permutation is possible in the main clause of a variety of sentence types, including passives. The change of word order does not affect the truth condition of the sentence, and therefore the V-initial counterparts have the same meaning as the V-final version.

(6)  a) Active sentence: V-final

  Zou-ga  tora-o   ker-ta-yo
  Elephant-NOM  tiger-ACC  kick-PAST-PART

  ‘The elephant kicked the tiger.’
b) Active sentence: V-initial
Ker-ta-yo zou-ga tora-o
Kick-PAST-PART elephant-NOM tiger-ACC
‘The elephant kicked the tiger.’

c) Passive sentence: V-final
Tora-ga zou-ni ker-are-ta-yo
Tiger-NOM elephant-by kick-PASS-PAST-PART
‘The tiger was kicked by the elephant.’

d) Passive sentence: V-initial
Ker-are-ta-yo tora-ga zou-ni
Kick-PASS-PAST-PART tiger-NOM elephant-by
‘The tiger was kicked by the elephant.’

Previous research on the acquisition of Japanese has pointed out that young children seem to be fairly familiar with the V-initial sentences. They appear in the spontaneous speech of 2-year-olds (Dansako, 2018), and when children produce a sentence in the V-initial order, they use the sentence in an adult-like manner (Dansako, 2018; Nomura, 2007; Sugisaki, 2008). These observations suggest that this word order is acquired quite early, and children at age 4-5 (the age range that is often associated with the passive difficulty) are expected to have a full command of the structure.

In relation to the IPH, the crucial property of V-initial passive sentences such as (6d) is that the passive morpheme -(r)are, a virtually unambiguous cue for the passive structure, appears before any overt NPs. Therefore, an incremental parser can determine the sentence is passive right at the beginning of the sentence, before encountering any NPs which it may map to thematic roles. Therefore, processing of V-initial passives such as (6d) does not require children to revise the initial misidentification of theta-roles, which is assumed to be the source of children’s difficulty with passives under the IPH. Consequently, the IPH predicts that children’s problem with passives diminishes in the V-initial order. If the V-initial order does not incur any independent difficulty for children, then children should simply show adult-like performance with V-initial passives. We carried out an experiment to test this prediction.

4. Experiment
4.1. Participants

The participants were 31 Japanese monolingual children at age 4;8-5;10 (Mean: 5;3). All of the children lived in central Tokyo areas, and attended a local private kindergarten. None were reported to have any language or developmental difficulties. Their caretakers signed a letter to approve the child’s participation to the experiment.
4.2. Design and materials

The experiment employed a 2x2 factorial design, with both WordOrder (V-final vs. V-initial) and Voice (Active vs. Passive) being within-participant factors. A sample set of crucial test sentences (=6)) are repeated here in (7).

(7) a) V-final / Active
   Zou-ga       tora-o   ker-ta-yo
   Elephant-NOM tiger-ACC  kick-PAST-PART
   ‘The elephant kicked the tiger.’

b) V-initial / Active
   Ker-ta-yo    zou-ga    tora-o
   Kick-PAST-PART elephant-NOM tiger-ACC
   ‘The elephant kicked the tiger.’

c) V-final / Passive
   Tora-ga      zou-ni    ker-are-ta-yo
   Tiger-NOM    elephant-by kick-PASS-PAST-PART
   ‘The tiger was kicked by the elephant.’

d) V-initial / Passive
   Ker-are-ta-yo   tora-ga   zou-ni
   Kick-PASS-PAST-PART tiger-NOM  elephant-by
   ‘The tiger was kicked by the elephant.’

Notice that in all of the four sentences, a declarative sentence-final particle -yo is attached to the verb. This was required to avoid a potential local ambiguity with the V-initial order. Since a relative clause in Japanese appears before the head noun, it is possible for the verb embedded in a relative clause to appear in the sentence-initial position, as shown in (8).

(8) Ker-ta  zou-ga  tora-o   mi-ta
    Kick-PAST  elephant-NOM Tiger-ACC   see-PAST
    ‘The elephant who kicked someone saw the tiger’

However, the verb within a relative clause cannot take the particle -yo, because the sentence-final particle can only be attached to the matrix verb:

(9) *Ker-ta-yo  zou-ga   tora-o   mi-ta
    Kick-PASS-PAST-PART elephant-NOM Tiger-ACC   see-PAST

Therefore, the presence of -yo in the test sentences (6b) and (6d) unambiguously rules out relative clause structure at the initial region of the sentence.
We prepared eight transitive verbs: osu ‘push’, hipparu ‘pull’, arau ‘wash’, tutuku ‘poke’, hikkaku ‘scratch’, keru ‘kick’, naderu ‘pat’, and kusuguru ‘tickle’. In a pilot experiment, we confirmed that Japanese children at around age 4-5 were familiar with these verbs and had no problem in understanding them. Four types of test sentences (as in (7)) were constructed for each of the 8 verbs, resulting in 32 sentences in total. The crucial test sentences, along with 4 practice sentences, were recorded by a female native Japanese speaker with Tokyo accent.

Along with the test sentences, we prepared cartoon pictures, each depicting a transitive event that corresponds to one of the 8 verbs. In each picture, two anthropomorphic characters participated in an event: one of them was the agent and the other was the patient (e.g., one pushing the other). There were 20 different characters (horse, dog, fox, racoon, lion, panda, gorilla, pig, zebra, rabbit, frog, bear, elephant, giraffe, cat, cow, tiger, sheep, koala, and monkey), and a total of 36 non-overlapping combinations of them were used in 32 test pictures and 4 practice pictures. In the pictures, the sides in which the agent appears were counterbalanced. The pictures were color-printed on A4-sized paper.

We created 4 experimental lists in which the 32 sentence-picture pairs were pseudo-randomized. Each list consisted of the 32 sentence-picture pairs, which included the same number of matching and mismatching conditions. The numbers of matching and mismatching conditions were counterbalanced within each experimental condition, and each participant was assigned one of the 4 lists.

4.3. Procedure

The task was picture-verification: the participant was presented with a picture-sentence pair, and was asked to judge whether or not the test sentence accurately described the picture. Each trial was organized as follows. First, the experimenter showed a picture to the participant, and asked the participant to name the characters on the picture. Then, the experimenter played a recorded test-sentence through a speaker built in a toy monkey (the child was told that the monkey was trying to describe the picture). The participant then judged whether the sentence matched the picture or not by pushing the “correct” or “wrong” toy button.

An experimental session began by 4 practice trials, with 2 active sentences and 2 passive sentences. After the practice, the test session with 32 trials commenced. During the test session, there were 2 breaks, in which the child participant was allowed to put a sticker on her personal progress chart. These interventions worked very well, keeping the child motivated and concentrated on the task when the trials resumed.

In the present study, we chose to employ picture-verification task over the standard Truth-Value Judgment Task (TVJT). This is because we put a priority on collecting enough data to carry out reliable within-participant comparisons, which required us to administer 8x4 test trials per child. A standard TVJT experiment with a story for each test sentence would take too much time for 4- and 5-year-olds to complete a session with 32 trials. Our picture-verification task
addressed this issue at the expense of eliminating experimental stories that build up a context for introducing the test sentences. It is possible that the lack of contexts in our experiment adversely affected children’s overall performance, but even if that was the case, we should still be able to gauge the effect of word order (V-final vs. V-initial) within the current design.

4.4. Predictions

The IPH predicts that, since a revision of theta-misassignment is not required with V-initial passives, the V-initial order should reduce children’s difficulty with passives. Thus, the most straightforward support for the IPH would be results that involve a significant effect of WordOrder and near-ceiling performance with V-initial passives. However, it is possible that the V-initial order induces an independent processing cost and then adversely affect children’s performance. If that is the case, then the negative effect of the V-initial order may cancel out the positive effect of the order for processing passives, resulting in a lack of improved performance with V-initial passives. In that case, we should find degraded performance with V-initial active sentences, and a significant interaction between the two experimental factors: WordOrder and Voice. Such a result would still be interpreted as evidence supporting the IPH.

4.5. Results

Figure 1 summarizes the overall results. As we can see in the graph, the correct response rates in the two Active conditions were very similar: 88.7% in V-final condition, and 85.8% in V-initial condition. The passive conditions show a similar trend: the correct response rates were 63.7% in V-final condition, and 65.3% in V-initial condition. We analyzed the number of correct responses with Linear Mixed-Effects models with a binomial function and selected the best-fit model using a backward selection approach (Jaeger, 2008). Voice (Active/Passive) and WordOrder (V-final/V-initial) were entered as fixed effects, and Participants and Items were random effects. The analysis revealed significant differences between Active and Passive in both V-final and V-initial conditions ($\beta=1.33, SE=0.16, z=8.1, p<0.001$). Thus, children performed better with active sentences than with passive sentences. This is a replication of the observation made in various previous studies. On the other hand, the effect of WordOrder was not significant ($\beta=0.09, SE=0.16, z=0.5, p=0.57$), nor was the interaction between Voice and Word-order ($\beta=0.32, SE=0.32, z=0.98, p=0.32$). That is, children’s performance with passives did not improve with V-initial sentences.
Figures 2a and 2b plot the distributions of individual response patterns with active test sentences (Figure 2a) and passive test sentences (Figure 2b). First, as shown in Figure 2a, most children were able to provide more than 6 correct responses out of 8 trials in the Active conditions. These results showed that children had no difficulty with actives.

On the other hand, as shown in Figure 2b, individuals are rather evenly distributed with respect to the number of correct responses within Passive conditions. This lack of bimodal distribution suggests that the development of passive comprehension does not involve an abrupt change at a certain age. In other words,
Japanese children’s passive comprehension seems to improve gradually as they get older.

We carried out an additional analysis to see whether the truth of the test sentence (True vs. False) affected children’s performance. We found a significant main effect of Truth ($p<0.01$), meaning that children performed better in the trials where the test sentence matched the picture. This effect is interpreted as reflecting children’s yes-bias. In addition, we found no significant interaction between Truth and Voice ($p=0.54$), or Truth and WordOrder ($p=0.69$). This suggests that the effect of children’s yes-bias was uniform across our experimental conditions, and does not pose any complications for the interpretation of the data.

5. Discussion

In the present study, we carried out an experiment with Japanese-speaking children to determine whether children’s problem with full passive comprehension could all be reduced to the general properties of their sentence processing mechanism, and whether we can find direct supporting evidence for the IPH. Under the IPH, the difficulty with full passive comprehension is reduced to the difficulty with revisions of initial processing commitments. Therefore, if no revision is required, no difficulty should be observed.

Our experiment utilized Japanese V-initial sentences, in which an unambiguous cue for passive structure (i.e., the passive morpheme -(r)are) appears in the initial position of the sentence. This word order should allow the parser to realize that it is dealing with a passive sentence before encountering any
NPs to which it might incrementally assign a thematic role. Therefore, no revisions of misidentified theta-roles should be required in the processing V-initial passive sentences. We sought to find out if the V-initial order help Japanese children to comprehend passive sentences. If children’s performance with passives improves with the V-initial order, then that would be regarded as direct supporting evidence of the IPH.

The results of our experiment, however, showed no significant effect of different word orders (V-final vs. V-initial), nor interaction between different voices and word orders. That is, we failed to find any evidence that suggests that the V-initial order made the processing of passive sentences easier for Japanese children. The results do not support the prediction from the IPH, and consequently, they do not allow us to conclude that the children’s difficulty with passives could be explained away in terms of general properties of their processing mechanism.

Before concluding, we must consider one potential interfering factor in our experiment: our V-initial test sentences might involve temporal ambiguity (H. Ono, personal communication). Compare the two sentences in (10): one is a single-clause sentence with V-initial order ((10a)), and the other involves a postposed adverbial clause ((10b)).

(10) a) Single clause structure
Ker-ta-yo zou-ga tora-o
Kick-PAST-PART elephant-NOM tiger-ACC
‘The elephant kicked tiger.’

b) Bi-clausal structure
Ker-ta-yo zou-ga tora-o mi-ta atoni
Kick-PAST-PART elephant-NOM tiger-ACC see-PAST after
Lit: ‘Kicked, after the elephant saw the tiger’

In (10a), the two NPs zou ‘elephant’ and tora ‘tiger’ are arguments of the verb keru ‘kick’. In (10b), in contrast, the two NPs belong to the subordinate adverbial clause, being arguments of the verb miru ‘see’. Since Japanese allows free omission of argument NPs, the arguments of the matrix verb keru ‘kick’ are construed as being dropped, and can refer to anyone appropriate in the context. Given that the verb appears in the sentence final position in the canonical word order in Japanese, an incremental parser may place a clausal boundary after the first verb ker-ta-yo, analyzing following NPs as belonging to a subordinate clause. Let us call this bi-clausal analysis.

(11) Incremental bi-clausal analysis
\[ [s \phi \phi \text{ker-ta-yo}] [s zou-ga ...] \]

If an incremental parser commits to such a bi-clausal analysis in processing the V-initial test sentences in our experiment, then it must perform a revision after realizing that there was no main predicate for the second clause. Now remember
that in our experiment, children did not perform better with V-final active sentences than with V-initial active sentences. This could mean either one of the followings: (i) children did not commit to the bi-clausal analysis, and therefore there was no revision, or (ii) children did commit to a bi-clausal analysis, but the revision cost associated with a V-initial active sentence was too small to affect the performance of 4- and 5-year-olds.

In our experiment, all the pictures depicted a simple transitive event with just two participants. Moreover, all the 32 test sentences were single-clause sentences. Thus, if children’s incremental parser took these into account, then it is possible that the parser simply regarded the bi-clausal analysis as implausible, and dismissed it before committing to the analysis. On the other hand, it is also possible that some types of revision are less problematic for children than others. For example, Omaki et al. (2014) found that Japanese-speaking children could successfully revise their initial parsing commitment if they were provided with a specific kind of revision cue. Thus, the lack of decrement of the correct responses in Active + V-initial condition does not necessarily lead to the conclusion that there were no revisions; children could have successfully revised their initial bi-clausal analysis with active test sentences.

An important point to note here is that revision costs of the bi-clausal analysis could be different for active and passive sentences. If the parser incrementally commits to the bi-clausal analysis in (11), then it would assume the predicate of the second clause is active rather than passive, and accordingly maps the NPs to a thematic role (i.e., the -ga marked NP to Agent, the -o marked NP to Patient, and the -ni marked NP to Goal). If that is the case, a revision into a single-clause active structure would not require revisions of theta-roles; but such thematic revisions are necessary to re-link the NPs to a passive predicate. Under this scenario, V-initial active sentences and V-initial passive sentences are different with respect to the nature of the revision required: only the latter involves revisions of theta-role assignments. And if that is the type of revision that children have problems with, then our data is still compatible with the IPH: our crucial test sentences (i.e., V-initial passives) actually required children to revise their initial parsing commitments in a way that led to degraded performance.

Taken together, our data is still open to several alternative interpretations. In order to test the bi-clausal analysis scenario, it is crucial to test if Japanese children actually commit to the bi-clausal analysis in the processing of V-initial sentences. It is also important to accumulate empirical data regarding relative difficulties with different types of revision. These are issues for future research.

References


