C<small>AMBRIDGE</small>
U<small>NIVERSITY</small> P<small>RESS</small>

ARTICLE

Children's online use of word order and morphosyntactic markers in Tagalog thematic role assignment: an eye-tracking study

Rowena GARCIA[1,2]* [iD], Jens ROESER[3], and Barbara HÖHLE[2,4]

[1]International Doctorate for Experimental Approaches to Language and Brain (IDEALAB): University of Potsdam, Germany; University of Groningen, the Netherlands; University of Newcastle, United Kingdom; University of Trento, Italy; and Macquarie University, Sydney, Australia, [2]University of Potsdam, Germany, [3]Nottingham Trent University, United Kingdom, and [4]Macquarie University, Sydney, Australia
*Corresponding author: Language Development Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands. E-mail: rowena.garcia@mpi.nl

Abstract

We investigated whether Tagalog-speaking children incrementally interpret the first noun as the agent, even if verbal and nominal markers for assigning thematic roles are given early in Tagalog sentences. We asked five- and seven-year-old children and adult controls to select which of two pictures of reversible actions matched the sentence they heard, while their looks to the pictures were tracked. Accuracy and eye-tracking data showed that agent-initial sentences were easier to comprehend than patient-initial sentences, but the effect of word order was modulated by voice. Moreover, our eye-tracking data provided evidence that, by the first noun phrase, seven-year-old children looked more to the target in the agent-initial compared to the patient-initial conditions, but this word order advantage was no longer observed by the second noun phrase. The findings support language processing and acquisition models which emphasize the role of frequency in developing heuristic strategies (e.g., Chang, Dell, & Bock, 2006).

Keywords: Tagalog acquisition; sentence comprehension; word order; morphosyntax; thematic role assignment

Introduction

In daily communications, we often have to identify the agent and the patient of an action described in a sentence that we hear. Therefore, it is crucial in language acquisition for children to learn how their language marks these agent and patient thematic roles, and to integrate this knowledge in their sentence processing. Moreover, they have to do this role assignment rapidly in the ongoing process of sentence interpretation. Identifying the strategies that children use to perform this task is crucial in deepening our understanding of language acquisition and

processing. The current study investigates thematic role assignment in children learning Tagalog – a language that has a complex but reliable system of morphosyntactic markers of thematic roles.

Previous research has shown that thematic role assignment can be a challenge for children's sentence comprehension, especially for sentences with non-canonical argument order (patient-before-agent, from here on referred to as non-canonical sentences) such as passives (Armon-Lotem et al., 2016, for Catalan, Lithuanian, and Hebrew; Bever, 1970, for English; de Barros Pereira Rubin, 2009, for Portuguese; Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008, for German; Frankel, Amir, Frenkel, & Arbel, 1980, also for Hebrew; Hakuta, 1977, for Japanese; MacWhinney, Pleh, & Bates, 1985, for Hungarian). Children tend to incorrectly interpret the first noun phrase (NP1) as the agent, thus reversing the thematic role assignments. This type of error in non-canonical sentence interpretation shows children's reliance on word order, which has been claimed to be due to the high frequency of sentences with an agent-before-patient order in the input (Demuth, 1989; Gordon & Chafetz, 1990; Kline & Demuth, 2010), and to the high reliability of this cue for assigning thematic roles in many languages (MacWhinney, 1987; MacWhinney & Bates, 1989). It is notable that this difficulty has been observed not only in languages with fixed word orders such as English, but also in languages with more flexible orders such as German.

Cross-linguistic differences in the use of a word order strategy have also been found, especially because some languages use other features such as case marking as cues to thematic role assignment. Previous studies suggest that children learning these languages begin to show a higher reliance on the morphosyntactic markers than on word order early on. For example, children speaking Serbo-Croatian begin to consistently rely on case marking at around four years of age, while Turkish-speaking children use case markers as early as two years (Slobin & Bever, 1982).

Recently, researchers investigated not only children's final sentence interpretation, but also how the sentence interpretation unfolds over time. Looking at real-time processing provides insights on whether children are 'only' slower in processing compared to adults, or if they use different strategies to arrive at an interpretation (Snedeker, 2013). Studies with adults have shown that they process incoming information in an incremental fashion and that information is not buffered until the end of a larger linguistic unit – such as the end of a sentence – before interpretation starts (Altmann & Steedman, 1988; Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003; Marslen-Wilson & Tyler, 1987).

Similar to adults, children have also shown evidence of incremental processing. In a seminal study, Trueswell, Sekerina, Hill, and Logrip (1999) found that the temporarily ambiguous phrase *on the napkin* in sentences such as "Put the frog on the napkin in the box" was initially interpreted as the goal of the action by adults and five-year-olds. However, unlike the adults, children did not revise their interpretation once the disambiguating phrase *in the box* was presented. The authors concluded that children process sentences incrementally akin to adults, but they have difficulty revising initial parses if these turn out to be inconsistent with the rest of the sentence. Other studies have also shown that children can incrementally use lexical information (Snedeker & Trueswell, 2004) and prosody (Snedeker & Yuan, 2008) in ambiguity resolution.

Children's strategy in processing sentences can be explained by models which emphasize the importance of frequency in forming comprehension strategies, such as Chang, Dell, and Bock's (2006) computational account of incremental word prediction and learning. According to this model, the parser continuously predicts the upcoming input from the previous input, but also through the use of event-semantic representations, including thematic roles. Therefore, an online interpretation of the first noun as the agent is automatically pursued if an agent-before-patient order is highly frequent in the input. For example, because English has a strong agent-before-patient bias, the model sets a strong weight for mapping the agent role to the first noun early in development. Through encountering deviations from this expected mapping of word order and thematic roles, the model gradually learns to put more weight on the post-first-noun structures for thematic role assignment.

To date, only a few experimental studies have focused on the real-time processing of non-canonical sentences in child language (Abbot-Smith, Chang, Rowland, Ferguson, & Pine, 2017; Huang, Zheng, Meng, & Snedeker, 2013; Schipke, Knoll, Friederici, & Oberecker, 2012; Zhou & Ma, 2018). These studies were interested in the timecourse of children's use of linguistic information such as word order and morphosyntax for thematic role assignment. For example, Schipke *et al.* (2012) showed using event-related potentials (ERP) that six-year-old German-speaking children processed accusative-marked nouns in the sentence-initial position similar to adults (same ERP patterns at the NP1), but three-year-olds did not show the same sensitivity to the case markers. However, while ERP studies provide evidence of children's sensitivity to a cue, it cannot clearly show children's interpretation of a sentence in real time.

Studies using eye-tracking provide more information on how children interpret sentences as they unfold. For example, Abbot-Smith and colleagues (2017) showed that English-speaking children (aged 2;1 to 3;5) incrementally map the first noun to the agent role, as the children were found to consistently look more to the clip which showed the first noun as the agent once they heard the initial noun of the sentence. Only the three-year-old children were able to revise their initial interpretation when the sentence turned out to be a passive sentence: after hearing the second noun phrase (NP2), the children in the passive condition showed fewer looks to the picture which showed the NP1 as the agent, in comparison to the children in the active condition.

Eye-tracking studies in Mandarin have also shown that three- and five-year-old children can rapidly use the voice markers BA (the noun phrase following this marker is assigned the patient role) and BEI (the noun phrase following this marker is designated as the agent) for sentence interpretation (Huang *et al.*, 2013; Zhou & Ma, 2018). In Zhou and Ma's study, children heard sentences like (1) which always had the word order BA/BEI Marker + Noun + Adverb + Verb. The other argument was dropped. Upon encountering the noun (*lion* in example 1), children already directed their gaze to the picture that showed the referent of the noun as the patient of the action in sentences that started with BA, and to the picture that showed the referent of the noun as the agent in sentences that started with BEI.

(1) BA/BEI   shizi   qingqingdi   bao-le   qilai
          lion   gently   hold   up
   'Someone gently holds / is held by the lion.'

Huang *et al.* (2013) used two arguments in their stimuli sentences and manipulated whether the NP1 was a noun or a pronoun. They asked five-year-old children to act out sentences like (2) "The seal is quickly eaten by it" while tracking their gaze pattern. Children were presented with three real objects at a time: (1) the mentioned item (seal); (2) a plausible agent of the action (shark); and (3) a plausible patient (fish). In addition to the incremental use of the morphosyntactic markers, the authors found that children were less likely to incorrectly interpret the first noun as the agent in the BEI condition when it was a pronoun ("It BEI seal quickly eat" or "It is quickly eaten by the seal") compared to when it was not ("Seal BEI it quickly eat" or "The seal is quickly eaten by it"). The authors proposed that the advantage of the pronoun condition indicates that children do not yet assign any role when they encounter a pronoun in the NP1 position, but that they only do so when the first noun is lexical. Therefore, in the former case, there is no need to revise an initial interpretation once the BEI marker is encountered, while the lexical noun condition requires a revision of the thematic role assignment.

(2) Haibao   BA/BEI   ta   henkuaijiu   chidiao
    Seal   it   quickly   eating
    'The seal is quickly eating it / eaten by it.'

These online studies have shown that children incrementally use different cues like word order and morphosyntactic markers depending on their age and target language. In Abbot-Smith *et al.*'s (2017) and Huang *et al.*'s (2013) experiments, the sentences were temporarily ambiguous when the NP1 occurred, as morphosyntactic markers were not yet present and thus could not play a role in interpretation. In Zhou and Ma's (2018) study, the initial argument was dropped, so there was no initial ambiguity. However, word order was still relevant for thematic role assignment in their study as the position of an argument relative to the markers needs to be considered for thematic role assignment in Mandarin.

There is evidence that children incrementally use word order for thematic role assignment but, in most of the previous studies, the morphosyntactic markers that were relevant for thematic role assignment (e.g., verb inflection and *by*-phrase in English, or the voice markers in Mandarin) only occurred after the first noun. Therefore, it is an open question whether morphosyntactic markers can lead to an immediate thematic role assignment to the first noun. In addition, it is an open question whether children incrementally assign the agent role to the first noun even in a language in which word order is structurally not relevant for thematic role assignment. In this paper, we investigated children's online comprehension of transitive sentences in Tagalog, a verb-initial language which does not use word order for thematic role assignment but instead uses verb and noun morphology that are given early on in the sentence, such that there is no initial ambiguity in interpretation.

In Tagalog, the verb carries voice, aspect, and mood information. The voice affix on the verb denotes the thematic role of the noun phrase that is marked by *ang* (from here on referred to as the *ang*-phrase) (Himmelmann, 2005).[1] In the agent voice (AV), the verb infix –*um*– assigns the *ang*-phrase the agent role (3, 5), while in the patient voice

---

[1]In Tagalog, voice-marking and mood are conflated. The voice-markings used in this paper also signal realis mood. See Himmelmann (2005) for a discussion of Tagalog voice-marking and mood.

(PV), the verb infix *–in–* assigns the *ang*-phrase the patient role (4, 6). Cooreman, Fox, and Givón (1984) found in a written corpus that, in sentences with transitive verbs, the patient voice occurs more frequently than the agent voice.

(3) H<**um**>ihila ng baboy **ang** baka
    <AV>[2]pull NSBJ pig SBJ cow
    'The cow is pulling a pig.'

(4) H<**in**>ihila ng baboy **ang** baka
    <PV>pull NSBJ pig SBJ cow
    'The/A pig is pulling the cow.'

(5) H<**um**>ihila **ang** baka ng baboy
    <AV>pull SBJ cow NSBJ pig
    'The cow is pulling a pig.'

(6) H<**in**>ihila **ang** baka ng baboy
    <PV>pull SBJ cow NSBJ pig
    'The/A pig is pulling the cow.'

The post-verbal argument order is relatively free (Schachter, 2015), and the basic order remains a matter of debate. From a grammatical perspective, word order is irrelevant for thematic role assignment in basic Tagalog sentences. Even if the order of the arguments differs between (3) and (5), the thematic roles are the same because both sentences are in the agent voice. Examples (4) and (6) also have the same meaning because they are in the patient voice.

Despite this flexible word order and the availability of morphosyntactic markers, Sauppe (2016) claimed, using evidence from a visual world paradigm eye-tracking experiment, that adult Tagalog speakers have a tendency to anticipate agent nouns to follow the verb (also providing evidence of incremental processing). He found that there were more looks to the agent image after the participants heard the verb, regardless of the voice-marking.

Tagalog-learning children's real-time processing of basic transitive sentences has not yet been deeply investigated, as most of the previous studies used offline measures. Segalowitz and Galang (1978) tested three-, five-, and seven-year-olds' comprehension of reversible transitive sentences with a sentence–picture matching task. Their results showed that children correctly interpreted patient voice agent-initial sentences (verb-agent-patient), but reversed the roles in agent voice patient-initial sentences (verb-patient-agent). Since voice and the order of arguments were confounded in this study, verb-medial sentences (agent-verb-patient for the agent voice and patient-verb-agent for the patient voice), which mostly occur in formal, written language, were used in a follow-up study. Children correctly interpreted both of these verb-medial sentences even if the patient voice was patient-initial, showing that they did not always use a word order strategy.

This patient voice advantage was also observed by Galang (1982). She claimed that children acquire the patient voice marker earlier than the agent voice inflection. She presented pictures with transitive actions (5 out of 15 were reversible) to three-, five-,

---

[2]The following abbreviations are used: AV for agent voice, PV for patient voice, LIN for linker, SBJ for subject, and NSBJ for non-subject.

seven-, and eight-year-old children, and instructed them with utterances like, *Ituro mo ang kum*akain (agent voice) / *ki*n*akain* (patient voice) 'Point to that which is eating/ being eaten'. Children of all age groups were more accurate in the patient voice compared to the agent voice, but results of statistical analysis were not reported.

Instead of a patient voice advantage, a study on relative clauses in Tagalog showed that five-year-old children performed better in interpreting agent relative clauses (verb is in the agent voice) than patient relative clauses (verb is in the patient voice) (Tanaka *et al.*, 2015). Since agent relative clauses were always agent-initial, and patient relative clauses were always patient-initial, the results also imply that children use a word order strategy in comprehending relative clauses.

A recent study by Garcia, Roeser, and Höhle (2019) used a combined self-paced listening and picture verification task to investigate Tagalog-learning children's use of word order and morphosyntactic markers for thematic role assignment. The results of their picture verification task showed that five- and seven-year-olds were more accurate in patient-initial sentences in the patient voice compared to the agent voice. In addition, adults and seven-year-olds showed longer listening times for the NP1 when the markers on the verb and the noun signaled a mismatch to the action in the picture, compared to when the markers and the picture matched. Meanwhile, the five-year-olds showed this effect only in the patient voice but not in the agent voice. The authors concluded that children relied more on the morphosyntactic markers in the patient voice, and on a word order strategy in the agent voice. They attributed the better performance in the patient voice to the high frequency of the patient voice in the input, which they found to comprise 53% of transitive verbs with at least one noun phrase in Marzan's (2013) child-directed speech corpus. On the other hand, agent voice-inflected verbs comprised only 21% of these utterances. Both voices were dominantly agent-initial (agent voice: 95%; patient voice: 85%).

These studies provide some evidence that Tagalog-learning children use an NP1-as-agent strategy, at least in the agent voice. However, the previous studies could not demonstrate how and when children assign thematic roles to the noun phrases that they encounter while the sentence unfolds. To close this gap, the current study combined a picture-selection task with eye-tracking, which allows for an online observation of the ongoing interpretation process.

Based on Chang *et al.*'s (2006) model, we would expect Tagalog-speaking children to incrementally interpret the NP1 as the agent, given that the input they receive mostly has an agent-before-patient order (Garcia *et al.*, 2019). Given that Chang *et al.*'s (2006) model uses error-based learning, the model also predicts that children would learn to use the morphosyntactic markers in the patient voice earlier than in the agent voice, given that they encounter more patient voice patient-initial sentences in their input than agent voice patient-initial sentences.

In the current experiment, children and adults were given a picture selection task while their looks to the screen were tracked. They first saw two pictures of a reversible action between two animals, e.g., 'a cow pulling a pig' and 'a pig pulling a cow'. They then heard a sentence describing one of the two pictures. Their task was to identify which picture matched the sentence. We crossed Voice (agent voice, patient voice) and Word Order (agent-initial, patient-initial) to create the experimental items.

If children generally rely on word order, and expect the NP1 to be the agent, they would show more looks to the target picture in the agent-initial condition compared to the patient-initial condition regardless of the voice-marking on the verb. In

contrast, if they can incrementally use the morphosyntactic markers for thematic role assignment, they are expected to start looking at the correct picture in all conditions when the NP1 of the sentence is presented. However, if eye-tracking data mirror the accuracy data, based on Garcia *et al.*'s (2019) results, we would expect children to incrementally use the morphosyntactic markers more efficiently in the patient voice than in the agent voice, resulting in more looks to the target in the patient voice patient-initial condition, compared to the agent voice patient-initial condition.

Method

Participants

Sixty-five children were recruited from Metro Manila, Philippines.[3] All children were from Tagalog-speaking households, and were reported to have Tagalog as their dominant language. Moreover, these children were selected based on teachers' reports that they had age-appropriate language and cognitive skills. The 33 five-year-old children (mean age: 5;4; age range: 5;0–5;10; males: 18) were kindergarten students from a public elementary school, while the 32 seven-year-olds (mean age: 7;4; age range: 7;0–7;10; males: 13) were Grade 2 students from the same school. Data from one five-year-old participant was excluded because of errors in the fillers.

We also recruited 32 adults from Metro Manila (mean age: 20; range: 18–27; males: 9). No participant was reported to have a history of language delay, or psychiatric or neurologic disorder. Adult participants and the parents of the children provided informed consent.

Materials

Sixteen transitive verbs (*hila* 'pull', *silip* 'peek at', *sipa* 'kick', *huli* 'capture', *palo* 'hit', *pasan* 'give a piggyback ride', *kagat* 'bite', *tira* 'hit', *sagip* 'rescue', *gamot* 'cure', *pili* 'choose', *tawag* 'call', *salo* 'catch', *karga* 'carry', *baril* 'shoot', and *habol* 'chase') were used to create semantically reversible sentences (also used in Garcia *et al.*, 2019). In such sentences, both of the nouns that fill the argument positions of the verb can be the agent or the patient of the verb, e.g., *Humihila ang baka ng baboy* 'The cow is pulling a pig'.

Voice (agent voice, patient voice) and Word Order (agent-initial, patient-initial) were crossed in the stimuli sentences, resulting in four conditions (Table 1). Animals served as the agents and patients in the sentences. A temporal adverb was placed after the NP1 to prolong the time before the NP2 was given, thus allowing more time to observe how the NP1 information is used for sentence interpretation. A spatial adverb was also added after the NP2 in order to have more time to observe the use of the NP2.

For each verb, target and distractor pictures (showing the reversal of the agent and the patient roles) were created (see Figure 1 for an example), resulting in 16 picture pairs. The side of the picture where the agent appeared and the direction of the action were counterbalanced. The target and distractor were surrounded either by a blue or a red frame. Each picture pair was used twice with only the assignment of

---

[3]The number of participants in this experiment is comparable to that of Zhou and Ma's (2018) study. Given this sample size, they found an effect of morphosyntactic marker type on three- and five-year-old children's looks to the target picture during online thematic role assignment.

Table 1. Sample experimental items for the verb *hila* 'pull', given a picture of a cow pulling a pig

| a	H<**um**>ihila	**ang**	baka	tuwing umaga	ng	baboy	sa maputik	na bukid.	agent voice agent-initial
	<AV>pull	SBJ	cow	every morning	NSBJ	pig	in muddy	LIN field	
b	H<**um**>ihila	ng	baboy	tuwing umaga	**ang**	baka	sa maputik	na bukid.	agent voice patient-initial
	<AV>pull	NSBJ	pig	every morning	SBJ	cow	in muddy	LIN field	
	'The cow is pulling a pig every morning in the muddy field.'								
c	H<**in**>ihila	ng	baboy	tuwing umaga	**ang**	baka	sa maputik	na bukid.	patient voice agent-initial
	<PV>pull	NSBJ	pig	every morning	SBJ	cow	in muddy	LIN field	
d	H<**in**>ihila	**ang**	baka	tuwing umaga	ng	baboy	sa maputik	na bukid.	patient voice patient-initial
	<PV>pull	SBJ	cow	every morning	NSBJ	pig	in muddy	LIN field	
	'The/A pig is pulling the cow every morning in the muddy field.'								
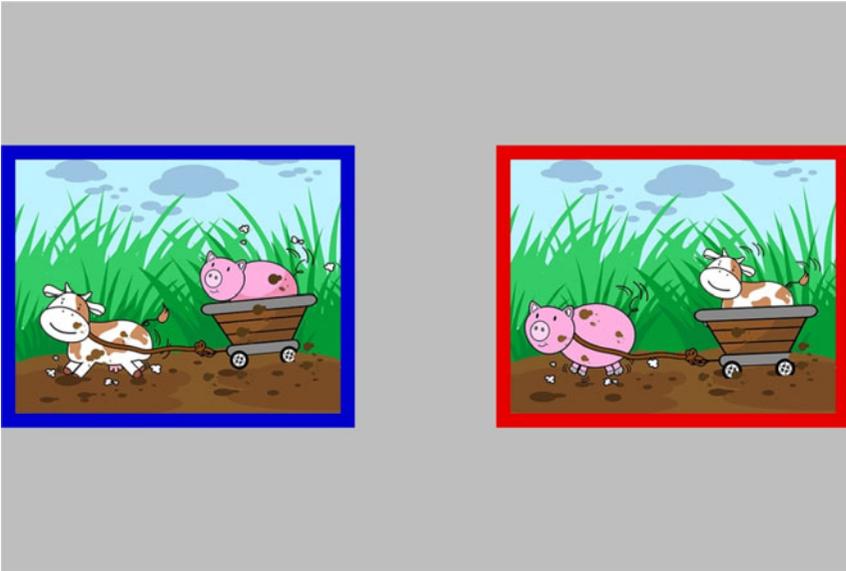
Figure 1. Sample picture pair for the experimental item "The cow is pulling a pig every morning in the muddy field".

the color of the frames to the two pictures being interchanged (resulting in 32 framed pairs). The side of the screen where the target picture appeared was also counterbalanced.

For the fillers, 16 other transitive verbs (e.g., *kain* 'eat', *inom* 'drink', and *basa* 'read') were used to create non-reversible sentences. They appeared in the same four conditions as the experimental items. The same animals were used as agents, while common inanimate concepts like *book*, *mango*, and *house* served as themes. Temporal and spatial adverbs were also used, so the sentence length matched that of the experimental items. Target and distractor (incorrect agent or theme) filler images were created.

The experimental and filler sentences were recorded in an audio booth by a native Tagalog speaker using a normal speaking rate. The Audacity 2.1.0 program (Audacity Team, 2015) was used for recording. The audio-recorded sentences (64) were combined with their corresponding picture pairs (note that each sentence and picture pair were used twice to control for the color of the frames, i.e., whether the blue frame was on the left or the right side of the screen) and then turned into a video using Adobe Flash CS3 Professional Version 9.0. The framed target and distractor pictures (460×356 pixels in size) appeared in the middle of the screen with a gray background. After 2000 ms from visual stimulus onset, the audio-recorded stimulus sentence started to play. The visual stimulus remained on the screen throughout the audio presentation, and for around 3000 ms after the end of the sentence. Each experimental item was 11000 ms long.

Each sentence–picture pairing (128) was distributed into eight different lists, following a Latin square design. In each list, each lexical verb and each target–distractor picture pair appeared only once. Each list contained 16 experimental items (4 from each condition) and 16 fillers. Half of the lists had the blue frame on the

left and the red frame on the right, while the other half had the red frame on the left and the blue frame on the right.

Procedure

Children were individually tested in quiet rooms in the schools, and the adult participants in a room at the university. The experimenter was seated next to each participant, and presented the experiment on a 17-inch laptop with a 1024×768 pixel resolution. An SMI RED-mobile eye-tracker with 60 Hz sampling rate was placed below the laptop's screen to record the participants' eye-movements. The stimuli were presented with SMI's Experiment Center 2 in a pseudo-randomized order such that no condition was presented more than two times in a row. The acoustic stimuli were presented through headphones. Each participant was tested with only one of the lists.

At the beginning of the experiment, the experimenter checked whether the children were familiar with the animals in the stimuli by asking them to point to the animal which she labeled, with four animals presented at a time. Children's knowledge of the verbs was also tested by asking them to point to the picture showing the action denoted by the uninflected verb that the experimenter said. These pictures showed two boys performing different actions instead of the animals used for the pictures in the main experiment. This task was also given to the adults just for consistency in data collection between the children and the control group. If a participant made an error during this pre-experiment phase, the experimenter gave reminders to look more closely at the pictures, and to listen more carefully. The experimenter proceeded to a five-point calibration of the eye-tracker if the participant successfully identified all the animals and verbs. Practice trials were given after the calibration phase.

In the practice trials, the participants were presented with items similar to the fillers used in the experiment (i.e., non-reversible actions). They first saw the target and distractor pictures, then heard the stimulus sentence. They were asked to name the color of the frame of the picture that matched the sentence they heard. Feedback was given during the practice trials. They were also verbally reminded not to point to the picture that matched the sentence. A verbal response was preferred to pointing, as pointing was expected to initiate larger movements. The experimenter proceeded to the experiment if the participant correctly answered at least three out of the four practice items.

In the experiment, the instructions were the same as in the practice trials but no feedback was given. The experimenter manually recorded the responses. A validation of the calibration was done after the last stimulus sentence was presented, in order to check whether the participant considerably moved from his/her position after the beginning of the experiment.

Data analysis

The experiment involved a $2 \times 2 \times 3$ factorial design. The independent variables used were Voice (agent voice and patient voice), Word Order (agent-initial and patient-initial), and Age Group (five-year-olds, seven-year-olds, and adults). The dependent variables were accuracy in the picture selection task, and the percentage of looks to the target picture (PLT).

Table 2. Average lengths of each time-window per condition in ms

	Verb	NP1	Temporal adverb	NP2	Spatial adverb
agent voice agent-initial	1099	934	1281	1040	884
agent voice patient-initial	1056	1081	1314	992	881
patient voice agent-initial	1065	1041	1343	921	870
patient voice patient-initial	1106	958	1265	1094	864

For the eye-tracking data, five time-windows were analyzed. The first time-window encompassed the verb (see Table 2 for the length of each window per condition). The second time-window corresponded to the NP1. The third time-window contained the temporal adverb. The fourth covered the NP2. The fifth time-window contained the first two words of the spatial adverb. Only the first two words of the spatial adverb were considered in order to make this time-window more similar in length to the other time-windows. The PLT was calculated by dividing the fixation on the target by the sum of fixations to the target and the distractor. These percentages were transformed into empirical logits for the statistical analyses.

R statistical software version 3.2.5 was used to perform the statistical analyses (R Core Team, 2016). Bayesian hierarchical models (Gelman *et al.*, 2014) were fitted using the rstanarm package (Stan Development Team, 2016). For both accuracy and PLT per time-window, the models were fitted with predictors for Voice, Word Order, and Age Group, two-way interactions of Voice and Word Order, Voice and Age Group (5:7, children:adults), and Word Order and Age group (5:7, children: adults), and three-way interactions of Age group, Word Order, and Voice. We used Helmert contrasts for the Age Groups (five-year-olds were compared to seven-year-olds, and both age groups of children were compared to the adults), and sum contrasts for Voice and Word Order. Random by-subject and by-item intercepts, and by-item slope adjustments, were included in the models. By-subject slopes were also included for the variables Voice and Word Order, and their interactions, but not for Age Group because the latter was a between-subjects factor.

We used weakly informative priors for the predictors in the models. From the posterior samples of the Bayesian model, we calculated the 95% uncertainty intervals (reported inside [ ] in this paper). Support for an effect on the dependent variable is indicated by uncertainty intervals that do not contain zero as a possible parameter value. The proportion of posterior samples smaller than 0 ($P(b < 0)$) was also calculated for each predictor. (It must be noted that $P(b < 0)$ is not the same as *p*-values in the frequentist approach.) This proportion indicates the probability of a negative effect, given the data. When $P(b < 0)$) approaches one, a negative effect (e.g., lower accuracy, fewer looks to the target) is supported. On the other hand, $P(b < 0)$ approaching zero supports a positive effect. $P(b < 0)$ values near 0.5 show no evidence for an effect. For a specific introduction to Bayesian statistics in psycholinguistics and developmental research, see Nicenboim and Vasishth (2016) and Van de Schoot *et al.* (2014).
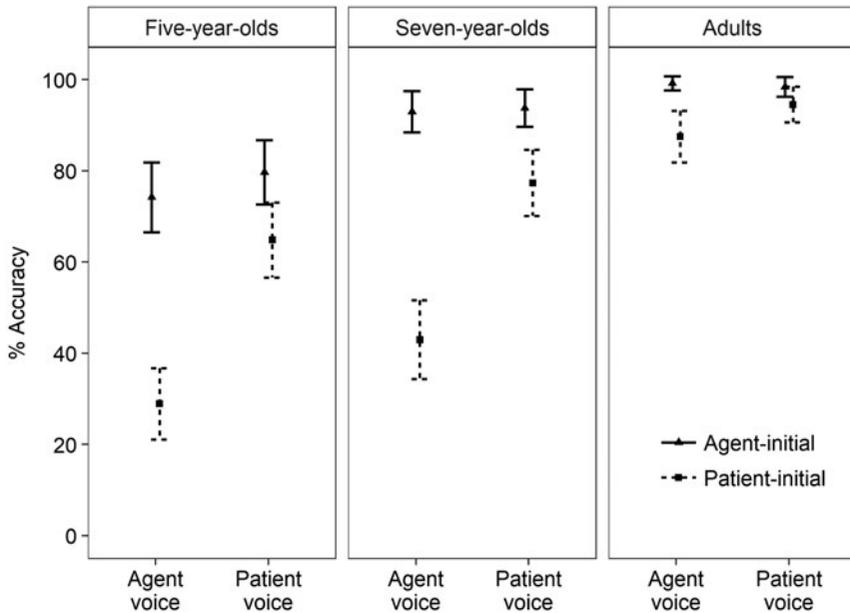
Figure 2. Mean accuracy with 95% confidence intervals for each condition per age group.

Results

We first present the accuracy data from the picture selection task, followed by the eye-tracking data.

Accuracy

The mean accuracy and 95% confidence intervals for each condition are shown in Figure 2. The Bayesian mixed effects model showed main effects of Age Group and Word Order; and two-way interactions of Age Group (5:7) and Word Order, and Voice and Word Order (see Table 3). Nested comparisons inspecting the two-way interaction of Age Group (5:7) and Word Order showed that the seven-year-olds scored higher than the five-year-olds in both Word Order conditions, but this difference was more pronounced in the agent-initial condition (coefficient = 3.60, [2.26, 5.00], $P(b < 0) < .001$) than in the patient-initial condition (coef = 1.45, [0.54, 2.38], $P(b < 0) = .001$). Nested comparisons inspecting the interaction of Voice and Word Order showed an agent-initial over patient-initial advantage for both voices, but with Word Order having a greater effect in the agent voice (coef = 9.82, [6.69, 13.59], $P(b < 0) < .001$) than in the patient voice (coef = 3.98, [1.20, 7.00], $P(b < 0)$ = .004). The Voice and Word Order interaction also showed higher accuracy in the patient voice compared to the agent voice in the patient-initial condition (coef = 4.57, [2.19, 6.93], $P(b < 0) < .001$) but not in the agent-initial condition (coef = −1.27, [−5.26, 2.44], $P(b < 0) = .75$). However, nested comparisons also showed that this effect of Voice in the patient-initial condition was found only in children (coef = 3.52, [2.00, 5.07], $P(b < 0) < .001$) and not in adults (coef = 1.05, [−0.17, 2.25], $P(b < 0) = .04$).

Table 3. Summary of the fixed effects in the Bayesian model of the participants' accuracy, including coefficients, 95% uncertainty intervals, and P($b < 0$), which refers to the probability that the true parameter value is less than 0

Comparison	Coefficient	Lower	Upper	P($b < 0$)
Intercept	2.20	1.82	2.64	< .001
Age Group(5:7)	5.04	3.40	6.80	< .001
Age Group(children:adults)	22.32	17.19	29.73	< .001
Voice	−3.37	−7.54	1.36	.93
Word Order	13.71	9.58	18.43	< .001
Age(5:7) Group*Voice	0.60	−1.04	2.27	.24
Age(5:7) Group*Word Order	2.14	0.56	3.79	< .001
Age Group(children:adults)*Voice	4.12	−1.34	10.88	.09
Age Group(children:adults)*Word Order	2.02	−3.44	9.17	.25
Voice*Word Order	5.76	1.40	10.56	.01
Age Group(5:7)*Voice*Word Order	0.59	−1.03	2.23	.24
Age Group(children:adults)*Voice*Word Order	1.25	−4.38	7.93	.33

To check for chance-level performance, we calculated 95% uncertainty intervals and the posterior probability that the accuracy was below chance (P($b < .5$)) from the posterior samples of the accuracy model. The uncertainty intervals are expected to contain the chance-level threshold (0.5) if responses are not different from chance. In the agent voice patient-initial condition, the five-year-olds scored below chance (coef = 0.27, [0.16, 0.40], P($b < .5$) > .99); the seven-year-olds scored at chance-level (coef = 0.43, [0.28, 0.58], P($b < .5$) = .83); and the adults above chance (coef = 0.90, [0.83, 0.96], P($b < .5$) < .001). All the other conditions were above chance for all age groups.

Eye-tracking data

We analyzed the proportion of looks to the target for each time-window: verb, NP1, temporal adverb, NP2, and spatial adverb (see Figure 3 for the adults, Figure 4 for the five-year-olds, and Figure 5 for the seven-year-olds). Each time-window was shifted by 200 ms to consider the time needed to program saccadic eye-movements (Matin, Shao, & Boff, 1993). Trials with more than 50% track loss in the time-window being analysed were excluded (0.01%). Moreover, we grouped the data into 250 ms time bins.

Figure 3 suggests that, immediately after the NP1, adults started looking to the target in all four conditions. Figure 4 indicates that five-year-olds' looks were still around chance-level after the NP1, but by the NP2 they already looked to the target in most of the conditions, except for the agent voice patient-initial condition where they showed more looks to the distractor. Figure 5 suggests that, at the temporal adverb time-window, seven-year-olds looked more to the target in the agent-initial conditions compared to the patient-initial conditions. However, by the NP2
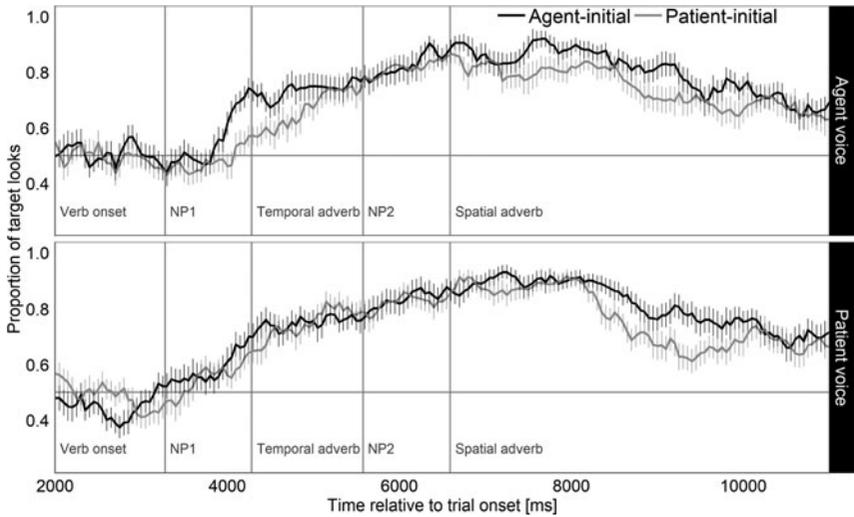
Figure 3. Adults' mean proportion of target looks with 95% confidence interval per condition relative to trial onset.
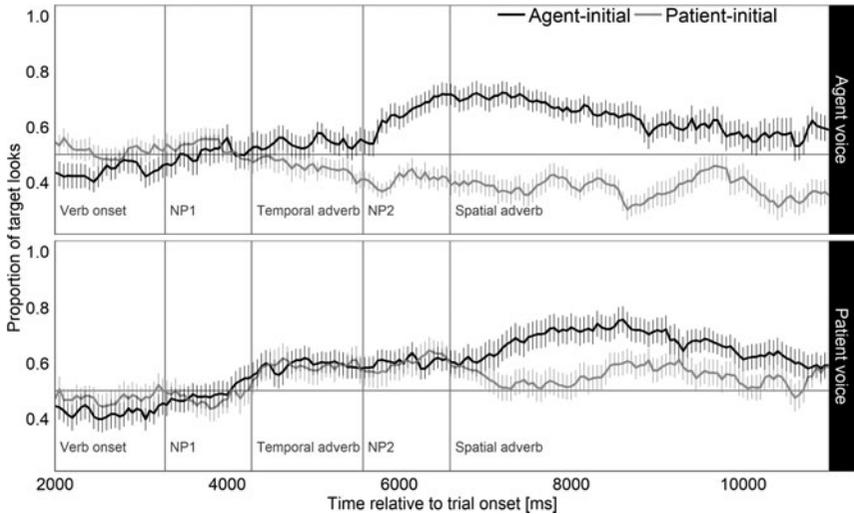

Figure 4. Five-year-olds' mean proportion of target looks with 95% confidence interval per condition relative to trial onset.

time-window, their PLT in the patient voice patient-initial condition has reached a similar level as in the agent-initial conditions, but has remained low in the agent voice patient-initial condition.

The Bayesian mixed model showed evidence for an effect of Age Group (5:7) on the PLT in the verb time-window, with seven-year-olds looking more to the target compared to the five-year-olds (see 'Appendix'). However, this time-window is prior to the disambiguation point which was the NP1, where we found no effect of the
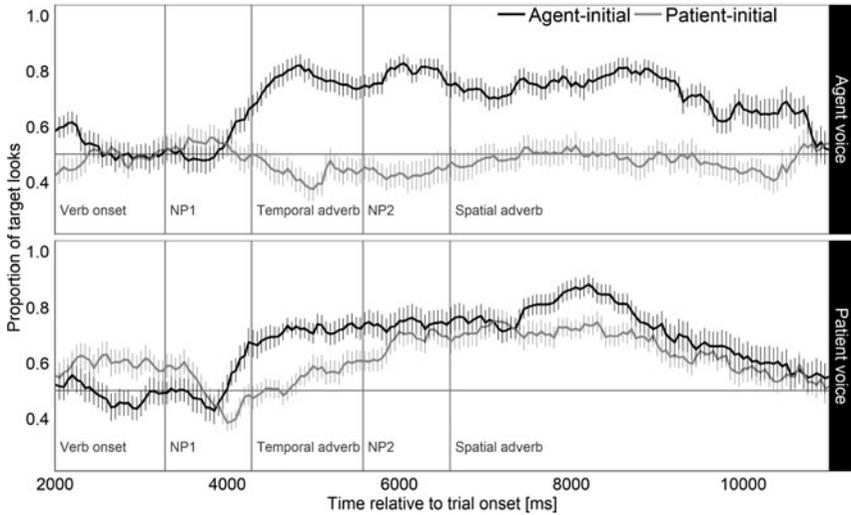
Figure 5. Seven-year-olds' mean proportion of target looks with 95% confidence interval per condition relative to trial onset.

independent variables on the PLT. In the succeeding temporal adverb time-window, there were main effects of Age Group (5:7 and children:adults) and Word Order; and two-way interactions of Age Group (5:7 and children:adults) and Word Order. Inspecting the interaction of Age Group and Word Order showed that children had higher PLT in the agent-initial compared to the patient-initial condition (coef = 3.70, [2.20, 5.23], P($b < 0$) < .001), while the adults did not show a difference between the Word Order conditions (coef = 0.39, [−0.61, 1.35], P($b < 0$) = .21). Checking the effect of Word Order in the two children groups separately showed that this effect of Word Order was present in the seven-year-olds (coef = 2.90, [1.93, 3.91], P($b < 0$) < .001), but not in the five-year-olds (coef = 0.8, [−0.22, 1.85], P($b < 0$) = .06).

In the NP2 time-window, the model showed main effects of Age Group, Voice, and Word Order, and two-way interactions of Age Group (children:adults) and Word Order, and Voice and Word Order, and a three-way interaction of Age Group (children:adults), Voice, and Word Order. Seven-year-olds had more looks to the target compared to the five-year-olds. Nested comparisons showed that children had higher PLT in the agent-initial compared to the patient-initial condition in the agent voice (coef = 3.51, [2.42, 4.59], P($b < 0$) < .001), but not in the patient voice (coef = 0.56, [−0.60, 1.69], P($b < 0$) = .16). However, there was no effect of Word Order found in the adults (agent voice: coef = 0.04, [−0.68, 0.75], P($b < 0$) = .46; patient voice: coef = −0.09, [-0.8, 0.59], P($b < 0$) = .59). The same results were obtained in the succeeding spatial adverb time window.

Discussion

We examined children's online use of word order and morphosyntactic markers for thematic role assignment in Tagalog. More specifically, we investigated whether Tagalog-speaking children incrementally interpret the NP1 as the agent of the

sentence, even if the verbal and nominal morphology to assign thematic roles are given early in Tagalog sentences. Incremental sentence interpretation was investigated by recording children's and adults' eye-movements to target and distractor pictures while listening to simple transitive sentences. The listening task was completed by a picture selection task.

The adults showed high accuracy scores in the picture selection task across all conditions. All age groups showed higher accuracy in identifying the picture in agent-initial compared to patient-initial sentences. In children, we also found an effect of voice in the accuracy scores of the patient-initial condition, with lower performance in the agent voice compared to the patient voice. Moreover, in the agent voice patient-initial condition, seven-year-olds scored at chance level, while the five-year-olds scored below chance. In the patient voice patient-initial condition, both five- and seven-year-olds scored above chance.

Regarding the eye-tracking data, adults showed an increasing proportion of looks to the target immediately after the point of disambiguation (NP1) in all of the conditions. In the temporal adverb time-window (immediately after the NP1), five-year-olds did not show a preference for the target nor the distractor picture in any of the conditions. On the other hand, seven-year-olds looked more to the target in the agent-initial than in the patient-initial conditions in this time-window. By the NP2, both five-year-olds and seven-year-olds showed more looks to the target in the agent-initial condition compared to the patient-initial condition, but this word order effect occurred only in the agent voice. In the patient voice, there was no effect of word order, and both groups of children showed more looks to the target than to the distractor.

Both the accuracy and eye-tracking data are consistent with previous findings that agent-initial sentences are easier to interpret than patient-initial sentences for both adults (Ferreira, 2003) and children (Armon-Lotem *et al.*, 2016; Dittmar *et al.*, 2008; MacWhinney *et al.*, 1985; Slobin & Bever, 1982). In the agent voice, above chance performance in the agent-initial condition, and below chance performance in the patient-initial condition indicate that the five-year-olds relied on word order and interpreted the NP1 as the agent of the action, resulting in thematic role-reversals in the patient-initial condition. Chance-level performance of the seven-year-olds in the agent voice patient-initial condition shows that they did not consistently rely on word order for thematic role assignment, but it also demonstrates that these older children still did not show adult-like use of the agent voice marker on the verb for assigning the thematic role to the *ang*-phrase. It is not so surprising that seven-year-old children still did not perform like adults, given previous findings that German-speaking children use case markers which clearly disambiguate thematic roles only after the age of five (Dittmar *et al.*, 2008).

Adults' immediate looks to the target after the NP1 shows their immediate use of the morphosyntactic markers on the verb and the noun for thematic role assignment. This incremental processing strategy is consistent with Sauppe's (2016) findings with Tagalog-speaking adults, and also with conclusions from studies on other languages (Altmann & Steedman, 1988; Kamide, Altman, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003).

Children also showed evidence of incremental processing. The agent-initial advantage in the seven-year-olds' looks to the target after the NP1 shows an early influence of word order in thematic role assignment. Even when the morphosyntactic markers were given early, the seven-year-olds showed that their interpretation was still affected by an NP1-as-agent expectation. The five-year-olds

showed a preference for one picture over the other only when they had encountered the NP2. This finding could indicate that the younger children are still slower in their general processing (Kail, 1991), but it could also mean that they wait for more morphosyntactic evidence before committing to a specific sentence interpretation compared to adults or older children.

The findings of the current study support models which predict an early NP1-as-agent bias if an agent-before-patient order is highly frequent in the input (e.g., Chang *et al.*, 2006; Yang, 2002). The dominance of this argument order was reported by Garcia *et al.* (2019) for child-directed speech in Tagalog for both the agent and patient voice. The current results are also in line with findings from languages where thematic role assignment is ambiguous by the NP1, which have shown that children immediately assign the agent role to the first noun they encounter (i.e., Abbot-Smith *et al.*, 2017, for English; Huang *et al.*, 2013, for Mandarin). Moreover, the present study extends this finding of an incremental use of word order to a language where morphosyntactic markers clearly assign the thematic roles from the start of the sentence.

The first-noun-phrase-as-agent strategy we found is also in line with Primus's (2006) proposed principle regarding the preference for an agent to occur before the patient. Primus claims that a thematically independent role such as that of an agent tends to precede a thematically dependent role such as that of a patient. A patient is considered to be thematically dependent on the agent, because there would be no patient had there been no agent acting on it. This agent-before-patient order has also been claimed to be preferred because it iconically reflects how an agent performs an action that affects a patient (Cohn & Paczynski, 2013; Kemmerer, 2012). It can also be said that an agent-initial order is preferred because an agent is more conceptually accessible than a patient (Bock & Warren, 1985).

However, children did not always rely on an NP1-as-agent strategy. Children's above chance accuracy in the patient voice patient-initial condition shows that they were able to use the patient voice marker on the verb to assign the thematic role to the *ang*-phrase. Additionally, in the patient voice, by the NP2, children showed more looks to the target than to the distractor in both the agent-initial and patient-initial condition, which means that they used the morphosyntactic markers for thematic role assignment. However, it must be noted that they used the morphosyntactic markers later than the adults, which may again be because children have slower processing speed (Kail, 1991).

This patient voice advantage is similar to findings by Segalowitz and Galang (1978) and Garcia *et al.* (2019) that Tagalog-speaking children rely more on word order than morphosyntactic markers in thematic role assignment in the agent voice, but rely more on the verb and noun morphology in the patient voice. Moreover, given that in Tagalog child-directed speech, there are more patient voice-marked verbs than agent voice-marked verbs (Garcia *et al.*, 2019), the patient voice advantage we found is comparable to findings of earlier acquisition of passives in languages where this structure is frequent (Demuth, 1989; Gordon & Chafetz, 1990; Kline & Demuth, 2010).

The patient voice advantage can also be explained by Chang *et al.*'s (2006) model. The model predicts error-based learning, which means that a structure is learned when an encountered word deviates from a predicted word. We can say that children first rely on a word order strategy because of its high frequency in the input, but this strategy can be overwritten by the morphosyntactic markers after a sufficient amount of patient-initial input was available. Given that the patient voice is more frequent,

and it is 85% agent-initial (compared to the agent voice's 95% agent-initial word order), there is a higher probability for children to encounter patient-initial sentences (a deviation from an NP1-as-agent expectation) in the patient voice compared to the agent voice, so children learn faster not to rely only on word order in the patient voice than in the agent voice.

As pointed out by an anonymous reviewer, the difference between the agent and the patient voice might also be due to the more restrictive interpretive possibilities for the *ng*-marked noun phrase of agent voice sentences. Although the *ng*-marked noun phrase of an agent voice sentence may be interpreted as specific or non-specific (Sabbagh, 2016), a non-specific interpretation is more common. However, our stimuli pictures require accessing the specific interpretation of the *ng*-noun phrase and therefore are in conflict with the more easily accessible non-specific interpretation of the *ng*-phrase in the agent voice. In contrast, the *ng*-phrase in the patient voice has no specificity constraint. This difference in specificity constraints could be the reason for the poorer performance in the agent compared to the patient voice. Then again, if the participants had issues with interpreting the *ng*-argument in the agent voice in our experiment, we would expect poor performance in the agent voice agent-initial condition and not only in the agent voice patient-initial condition. However, this prediction is not supported by our findings.

Our results also show that, in actual sentence parsing, word order appears to be used before the morphosyntactic markers. In the patient voice conditions where the seven-year-olds showed above chance accuracy, they initially showed more looks to the target in the agent-initial compared to the patient-initial condition, but this word order effect was no longer found in the later NP2 time-window. This finding implies that children have to first disregard an NP1-as-agent preference, before they could use the morphosyntactic markers. This provides some insights into why patient-initial sentences are generally more difficult to process than agent-initial sentences: children may have already interpreted the first noun as the agent before or upon encountering the NP1, but they were able to revise this interpretation once they have obtained more information from the morphosyntactic marker in the NP2. Such an explanation goes against previous findings in other languages that children have difficulties in revising an initial interpretation (Trueswell & Gleitman, 2004). Unfortunately, our paradigm does not allow us to clearly observe a revision, as it can also be that children just needed more time or cues to assign their initial interpretation.

Another limitation of the study comes from the lack of a comprehensive standardized assessment of the participants' language and cognitive skills. The children were chosen based only on teachers' reports that they had age-appropriate language and cognitive skills. However, the effects we have observed in the results were not only driven by the performance of a few participants – what we would expect if there were children with developmental disorders in our sample. Therefore, there is no reason to believe that the data found could be due to atypical language development of some of the children.

In conclusion, our study demonstrates that Tagalog-speaking children expect the first noun to be an agent, even though this language does not formally use word order for thematic role assignment. Although the morphosyntactic markers disambiguate the thematic roles before the nouns are given, it seems that children's word order expectation significantly affects sentence interpretation such that they need to hear more cues (NP2) to correctly assign the thematic roles in patient-initial sentences. These findings inform on the timing of use of the cues during sentence

interpretation, and show that research on under-studied languages can improve our understanding of language acquisition and processing.

References

Abbot-Smith, K., Chang, F., Rowland, C., Ferguson, H., & Pine, J. (2017). Two and three year old children use an incremental first-NP-as-agent bias to process transitive sentences. *PloS ONE*, *12*(10), 1–20.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191–238.

Armon-Lotem, S., Haman, E., Jensen de López, K., Smoczynska, M., Yatsushiro, K., Szczerbinski, M., … van der Lely, H. (2016). A large-scale cross-linguistic investigation of the acquisition of passive. *Language Acquisition*, *23*(1), 27–56.

Audacity Team (2015). Audacity(R): free audio editor and recorder (Version 2.1.0) [computer program]. Retrieved from<http://audacity.sourceforge.net/>.

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

Bock, J., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–72.

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: the processing of semantic roles in visual narrative. *Cognitive Psychology*, *67*(3), 73–97.

Cooreman, A., Fox, B., & Givón, T. (1984). The discourse definition of ergativity. *Studies in Language*, *8*(1), 1–34.

de Barros Pereira Rubin, M. C. (2009). The passive in 3- and 4-year-olds. *Journal of Psycholinguistic Research*, *38*(5), 435–46.

Demuth, K. (1989). Maturation and the acquisition of the Sesotho passive. *Linguistic Society of America*, *65*(1), 56–80.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development*, *79*(4), 1152–67.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164–203.

Frankel, D. G., Amir, M., Frenkel, E., & Arbel, T. (1980). A developmental study of the role of word order in comprehending Hebrew. *Journal of Experimental Child Psychology*, *29*(1), 23–35.

Galang, R. G. (1982). Acquisition of Tagalog verb morphology: linguistic and cognitive factors. *Philippine Journal of Linguistics*, *13*(2), 1–15.

Garcia, R., Roeser, J., & Höhle, B. (2019). Thematic role assignment in the L1 acquisition of Tagalog: use of word order and morphosyntactic markers. *Language Acquisition*, *26*(3), 235–61.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Chapman and Hall/CRC.

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, *36*, 227–54.

Hakuta, K. (1977). Word order and particles in the acquisition of Japanese. *Papers and Reports on Child Language Development*, *13*, 117–27.

Himmelmann, N. P. (2005). Tagalog. In A. Adelaar & N. P. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 350–76). London: Routledge.

Huang, Y. T., Zheng, X., Meng, X., & Snedeker, J. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language*, *69*(4), 589–606.

Kail, R. V. (1991). Development of processing speed in childhood and adolescence. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 25, pp. 151–85). New York: Academic Press.

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 (1), 133–56.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive procesing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32(1), 37–55.

Kemmerer, D. (2012). The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area. *Linguistics and Language Compass*, 6(1), 50–66.

Kline, M., & Demuth, K. (2010). Factors facilitating implicit learning: the case of the Sesotho passive. *Language Acquisition*, 17(4), 220–34.

MacWhinney, B. (1987). The Competition Model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B., & Bates, E. (Eds.) (1989). *The cross-linguistic study of sentence processing*. New York: Cambridge University Press.

MacWhinney, B., Pleh, C., & Bates, E. (1985). The development of sentence intrepretation in Hungarian. *Cognitive Psychology*, 17, 178–209.

Marslen-Wilson, W. D., & Tyler, L. (1987). Against modularity. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language processing* (pp. 57–104). Cambridge, MA: MIT Press.

Marzan, J. C. (2013). *Spoken language patterns of selected Filipino toddlers and pre-school children* (Unpublished doctoral dissertation). University of the Philippines.

Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: information processing time with and without saccades. *Perceptual Psychophysics*, 53(4), 372–80.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: foundational ideas–Part II. *Linguistics and Language Compass*, 10(11), 591–613.

Primus, B. (2006). Hierarchy mismatches and the dimensions of role semantics. In I. Bornkessel, M. Schlesewsky, B. Comrie, & A. Friederici (Eds.), *Semantic role universals and argument linking: theoretical, typological and psycholinguistic perspectives* (pp. 53–88). Berlin: Walter de Gruyter.

R Core Team (2016). R: a language and environment for statistical computing (Version 3.2.5) [computer program]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.

Sabbagh, J. (2016). Specificity and objecthood in Tagalog. *Journal of Linguistics*, 52(3), 639–88.

Sauppe, S. (2016). Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences. *Frontiers in Psychology*, 7, 1–13.

Schachter, P. (2015). Tagalog. In T. Kiss & A. Alexiadou (Eds.), *Syntax – theory and analysis: an international handbook* (Handbooks of Linguistics and Communication Science Vol. 42/3) (pp. 1658–76). Berlin: De Gruyter Mouton.

Schipke, C. S., Knoll, L. J., Friederici, A. D., & Oberecker, R. (2012). Preschool children's interpretation of object-initial sentences: neural correlates of their behavioral performance. *Developmental Science*, 15 (6), 762–74.

Segalowitz, N. S., & Galang, R. G. (1978). Agent-patient word-order preference in the acquisition of Tagalog. *Journal of Child Language*, 5(1), 47–64.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: a crosslinguistic study of word order and inflections. *Cognition*, 12, 229–65.

Snedeker, J. (2013). Children's sentence processing. In R. P. G. Van Gompel (Ed.), *Sentence processing* (pp. 189–220). New York: Psychology Press.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: the role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49 (3), 238–99.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, 58(2), 574–608.

Stan Development Team (2016). RStan: the R interface to Stan. R package version 2.14.1. Retrieved from <http://mc-stan.org/>.

Tanaka, N., O'Grady, W., Deen, K., Kim, C.-E., Hattori, R., Bondoc, I. P. M., & Soriano, J. U. (2015). Split focus preferences in Tagalog: evidence from child language. In A. Camp, Y. Otsuka, C. Stabile, & N. Tanaka (Eds.), *AFLA 21: The Proceedings of the 21st Annual Meeting of the Austronesian Formal Linguistics Association* (pp. 279–88). Canberra: Asia-Pacific Linguistics.

Trueswell, J. C., & Gleitman, L. (2004). Children's eye movements during listening: evidence for a constraint-based theory of parsing and word learning. In J. M. Henderson & F. Ferreira (Eds.), *Interface of language, vision, and action: eye movements and the visual world* (pp. 319–46). New York: Psychology Press.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, 73(2), 89–134.

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A gentle introduction to Bayesian analysis: applications to developmental research. *Child Development*, 85(3), 842–60.

Yang, C. (2002). *Knowledge and learning in natural language*. Oxford University Press.

Zhou, P., & Ma, W. (2018). Children's use of morphological cues in real-time event representation. *Journal of Psycholinguistic Research*, 47(1), 241–60.

Appendix

Summary of the fixed effects in the Bayesian model of the participants' looks to the target for each time-window, including Coefficients, 95% uncertainty intervals, and $P(b<0)$, which refers to the probability that the true parameter value is less than 0

Comparison	Coefficient	Lower	Upper	$P(b<0)$
Verb				
Intercept	−0.05	−0.19	0.11	.73
Age Group(5:7)	1.65	0.14	3.14	.02
Age Group(children:adults)	−0.79	−3.32	1.64	.74
Voice	0.32	−2.43	3.05	.40
Word Order	−1.65	−3.72	0.36	.95
Age(5:7)*Voice	−0.60	−1.85	0.81	.82
Age(5:7)*Word Order	0.36	−0.96	1.67	.30
Age Group(children:adults)*Voice	0.66	−2.91	4.12	.35
Age Group(children:adults)*Word Order	0.78	−2.46	4.16	.31
Voice*Word Order	1.07	−0.71	2.88	.12
Age Group(5:7)*Voice*Word Order	1.35	−0.53	3.23	.08
Age Group(children:adults)*Voice*Word Order	−0.57	−3.29	2.08	.66
NP1				
Intercept	0.14	−0.03	0.31	.04
Age Group(5:7)	0.42	−1.28	2.07	.30
Age Group(children:adults)	1.93	−0.70	4.72	.08
Voice	0.12	−2.43	2.69	.46
Word Order	0.74	−1.39	2.91	.24

(Continued)

(*Continued.*)

Comparison	Coefficient	Lower	Upper	P($b < 0$)
Age(5:7)*Voice	0.14	−1.65	1.84	.43
Age(5:7)*Word Order	0.12	−1.97	2.10	.46
Age Group(children:adults)*Voice	−1.81	−4.64	1.04	.90
Age Group(children:adults)*Word Order	1.60	−0.98	4.31	.11
Voice*Word Order	−0.19	−2.36	1.90	.56
Age Group(5:7)*Voice*Word Order	0.33	−1.28	1.92	.34
Age Group(children:adults)*Voice*Word Order	0.00	−2.40	2.24	.50
Temporal adverb				
Intercept	0.79	0.58	0.99	< .001
Age Group(5:7)	1.77	0.22	3.31	.01
Age Group(children:adults)	7.57	5.30	9.90	< .001
Voice	−1.80	−4.15	0.74	.92
Word Order	4.07	2.10	6.06	< .001
Age(5:7)*Voice	0.49	−1.03	2.00	.26
Age(5:7)*Word Order	2.12	0.75	3.42	< .001
Age Group(children:adults)*Voice	0.67	−1.58	2.85	.28
Age Group(children:adults)*Word Order	−2.93	−5.07	−0.82	0.99
Voice*Word Order	1.76	−0.18	3.68	0.04
Age Group(5:7)*Voice*Word Order	0.19	−1.19	1.63	.39
Age Group(children:adults)*Voice*Word Order	−0.90	−2.84	1.01	.84
NP2				
Intercept	1.08	0.84	1.32	< .001
Age Group(5:7)	2.12	0.43	3.87	.01
Age Group(children:adults)	10.57	8.25	13.06	< .001
Voice	−2.71	−4.94	−0.47	> .99
Word Order	4.06	1.68	6.18	< .001
Age(5:7)*Voice	−0.38	−1.93	1.07	.70
Age(5:7)*Word Order	1.40	−0.21	2.97	.04
Age Group(children:adults)*Voice	1.49	−0.79	3.73	.10
Age Group(children:adults)*Word Order	−4.18	−6.33	−1.98	> .99
Voice*Word Order	3.06	1.19	5.09	< .001
Age Group(5:7)*Voice*Word Order	0.36	−1.10	1.79	.31
Age Group(children:adults)*Voice*Word Order	−2.70	−4.71	−0.63	> .99

(*Continued*)

(*Continued.*)

Comparison	Coefficient	Lower	Upper	P($b < 0$)
Spatial adverb				
Intercept	1.18	1.00	1.37	< .001
Age Group(5:7)	2.33	0.71	4.09	< .001
Age Group(children:adults)	10.86	8.15	13.46	< .001
Voice	−2.85	−4.80	−0.86	> .99
Word Order	4.80	2.78	7.00	< .001
Age(5:7)*Voice	−1.08	−2.73	0.60	.90
Age(5:7)*Word Order	−0.17	−1.60	1.32	.59
Age Group(children:adults)*Voice	1.69	−0.55	3.96	.07
Age Group(children:adults)*Word Order	−2.74	−4.95	−0.53	.99
Voice*Word Order	2.94	1.03	4.76	< .001
Age Group(5:7)*Voice*Word Order	−0.37	−1.53	0.83	.72
Age Group(children:adults)*Voice*Word Order	−2.65	−4.51	−0.82	> .99