



Word order preferences of Tagalog-speaking adults and children

Journal:	<i>First Language</i>
Manuscript ID	Draft
Manuscript Type:	Original Manuscript
Keywords:	Tagalog acquisition, word order, child language acquisition, sentence production, voice
Abstract:	<p>We investigated the word order preferences of Tagalog-speaking adults and five- and seven-year-old children. The participants were asked to complete sentences to describe pictures depicting actions between two animate entities. Adults preferred agent-initial constructions in the patient voice but not in the agent voice, while the children produced mainly agent-initial constructions regardless of voice. This agent-initial preference, despite the lack of a close link between the agent and the subject in Tagalog, shows that this word order preference is not merely syntactically-driven (subject-initial preference). Additionally, the children's agent-initial preference in the agent voice, contrary to the adults' lack of preference, shows that children do not respect the subject-last principle of ordering Tagalog full noun phrases. These results suggest that language-specific optional features like a subject-last principle takes longer to be acquired.</p>

SCHOLARONE™
Manuscripts

1
2
3 A critical task in language acquisition is learning the specific word order regularities of
4 the ambient language. Children acquiring languages like English must learn that the
5 position of an argument in a sentence is crucial for determining who the agent and the
6 patient of an action denoted by the verb are. For example, to express that a girl named
7 Mary is kicking a boy named John, they should code Mary as the subject and therefore, at
8 the initial position of the sentence, and John as the object, which comes after the verb.
9
10

11 It is claimed that children generally acquire the language's canonical forms before
12 the non-canonical patterns, with *canonical* referring to structures which speakers produce
13 with minimal assumptions regarding the listener's background knowledge (Slobin, 1982;
14 Slobin & Bever, 1982). Studies on spontaneous speech have shown that children prefer a
15 subject-before-object word order in both fixed (Brown, 1973 for English; Slobin &
16 Bever, 1982 for Italian) and relatively more flexible word order languages (Lee, 2010 for
17 Mandarin; Slobin & Bever, 1982 for Serbo-Croatian, and Turkish; Tanaka & Shirai, 2012
18 for Japanese). The same preference was found in production experiments wherein
19 children had to describe pictures (Hakuta, 1982 for Japanese), videos (Cannizzaro, 2012
20 for Dutch and English), or act-outs (Angiolillo & Goldin-Meadow, 1982 for English,
21 Cannizzaro, 2012).
22
23
24

25 In most of these studied languages, the canonical order is subject-before-object,
26 and the agent usually corresponds to the subject (Dryer, 2013), which means that the
27 subject-before-object order is also agent-before-patient. This order reflects both the
28 grammatical relational hierarchy (the subject is higher than the object; Johnson, 1977)
29 and the thematic role hierarchy (the agent is higher than the patient; Fillmore, 1968;
30 Siewierska, 1993). The subject-before-object preference is considered to have such great
31 importance in word ordering patterns that Greenberg (1963) proposed it as Universal #1:
32 "In declarative sentences with nominal subject and object, the dominant order is almost
33 always one in which the subject precedes the object." Another – maybe not independent -
34 ordering principle relates to the thematic roles of the argument with the agent preferably
35 occurring before the patient. This order is considered to result from a universal principle
36 that the thematically independent role (agent) tends to precede and/or c-command the role
37 that is thematically dependent (patient) (Primus, 2006). A patient is thematically
38 dependent on the agent, because there would be no patient if there were no agent acting
39 on it in the first place. Primus (2003) proposed that this thematic dependency may be
40 derived from the dependency of an effect to a cause. Others have claimed that an agent-
41 before-patient preference reflect, in an iconic manner, how an agent initiates a causal
42 event which affects the patient (Cohn & Paczynski, 2013; Kemmerer, 2012).
43
44
45
46
47

48 In languages with a subject-before-object/agent-before-patient canonical order,
49 children do not face a conflict on which argument should occur first—it is the subject
50 which is also usually the agent, except for the passive voice. Therefore, in these
51 languages, it cannot easily be disentangled whether the choice of the word order is driven
52 by a subject-before-object preference or by an agent-before-patient preference. After all,
53 Jackendoff and Wittenberg (2014) have proposed that already at the two-word stage,
54 children have a preference for an agent-before-patient order in utterances containing only
55 nouns—an agent and a patient. They claim that this is a direct mapping from thematic
56 roles to linear position without the need to resort to grammatical relations, which means
57
58
59
60

that children prefer an agent to appear before a patient, even when the former is not a subject.

In this paper, we want to investigate children's acquisition of word order patterns in Tagalog – a language wherein the first noun phrase position is not confounded with the subject position, and the agent is not closely linked to the subject. These properties lead to the question of which word order Tagalog-learning children would prefer—subject-before-object or agent-before-patient. Studying the acquisition of Tagalog can thus show whether children's word order preference is syntactically-driven, i.e., subject-before-object, or semantically-driven (determined by thematic roles), i.e., agent-before-patient.

Tagalog voice-marking and word order

Tagalog, a Malayo-Polynesian Austronesian language with over 21 million speakers, is one of the major Philippine languages (Lewis, Simons, & Fennig, 2016). Most importantly for the current study, Tagalog is described as a symmetrical voice language, which means that the language has multiple basic transitive constructions, which are considered symmetrical because the verb bears a specific marker in all of the voice alternations (Himmelman 2005a; Riesberg, 2014). Furthermore, there is no demotion of an argument to an oblique sentence element across the voice alternations which is different from the active-passive alternation in other languages like English or German (Riesberg & Primus, 2015).

Tagalog's basic sentence structure includes the predicate and the so-called *ang*-phrase, which is the sentence subject (Guilfoyle, Hung & Travis, 1992; Kroeger, 1993a; see Schachter, 2015 for an alternative view). Other arguments as well as adjuncts may be preceded by the morphological marker *ng*, which can signify a common noun object, a possessor, or also an adjectival modifier (Schachter & Otanes, 1972), or by the morphological marker *sa*, which is a locative preposition (Himmelman, 2005b). The predicate is usually a verb which is inflected for mood, aspect, and voice.

In the Tagalog voice-marking system, a morphological marker on the verb assigns the thematic role of the *ang*-phrase or the subject (Himmelman, 2005b).¹ In the agent voice (AV), the verbal infix *-um-* (see Latrouite, 2001 for a discussion on affix choice), indicates that the *ang*-phrase is the agent, see example (1). In contrast, in the patient voice (PV), the verbal infix *-in-* denotes that the *ang*-phrase is the patient, see example (2). Hence, the roles of agent and patient are reversed with only a change in the voice-marking on the verb. The thematic role assignments are not affected by the order of the arguments such that there is no change in meaning between examples (1) and (3) in agent voice and between examples (2) and (4) in patient voice.

(1) H<um>ihila ng baboy **ang** baka
 AV-pull OBJ pig SUBJ cow
 'The cow is pulling a pig.'

(2) H<in>ihila ng baboy **ang** baka
 PV-pull OBJ pig SUBJ cow
 'The/A pig is pulling the cow.'

1
2
3
4
5 (3) H<um>ihila **ang** baka ng baboy
6 AV-pull SUBJ cow OBJ pig
7 'The cow is pulling a pig.'

8
9 (4) H<in>ihila **ang** baka ng baboy
10 PV-pull SUBJ cow OBJ pig
11 'The/A pig is pulling the cow.'

12
13
14 The choice of voice seems to be affected by several factors. Definiteness is one of
15 them (Himmelmann, 2005b): *ang*-phrases are always interpreted as definite, therefore
16 patient voice is used in sentences with definite patients. A written corpus study by
17 Cooreman, Fox, and Givon (1984), and a picture description task by Tanaka (2015)
18 provide empirical evidence that the patient voice is generally preferred when a patient is
19 present. However, Tanaka also showed that when the patient is inanimate, the patient
20 voice preference is weakened.

21
22
23 Another feature of Tagalog is its relatively free word order. The canonical order is
24 verb-initial, but the order of the arguments is not fixed (Schachter, 2015). An SVO order
25 is also grammatical, but it is considered more formal and is usually found in writing
26 (Schachter & Otnes, 1972). The basic orders of arguments and grammatical functions
27 are still matters of controversy. There are claims that the canonical order is VOS
28 (Billings, 2005), VSO (Aldridge, 2002), agent-initial (Buenaventura-Naylor, 1975;
29 Manuelli, 2010; Schachter, 2015), or VOS for the patient voice and both VOS and VSO
30 for the agent voice (Guilfoyle, Hung & Travis, 1992; Kroeger, 1993b).

31
32
33 In verb-initial sentences with a pronoun, the pronoun occurs immediately after the
34 verb (Billings, 2005). Concerning sentences with non-pronominal arguments, Kroeger
35 (1993b) proposed three principles which determine the preferred order of full noun
36 phrases. First, the agent tends to precede the other arguments (which we will call the
37 agent-first principle).² Second, the *ang*-phrase tends to be the last phrase (subject-last
38 principle). Third, heavier noun phrases (longer constituents) follow lighter noun phrases
39 (shorter constituents). The first and third principles are commonly observed across
40 languages but the second seems to go against the widely observed subject-before-object
41 preference.

42
43
44 There have been a few experimental studies that shed light on speakers'
45 preferences of ordering non-pronominal arguments in Tagalog. Manuelli (2010)
46 manipulated the voice-marking of the verb *kain* 'eat,' and the order of the arguments *fish*
47 and *cat*, and asked native adult speakers of Tagalog to rate the grammaticality of the
48 sentences, such as *Kumakain ng isda si Muning* ('The cat is eating a fish'). All 11
49 participants judged the agent voice patient-initial and patient voice agent-initial orders
50 (VOS) as grammatical, while 3 participants judged the agent voice patient-initial and
51 patient voice patient-initial (VSO) as less grammatical. More recently, Hsieh (2016) used
52 more verbs in a similar task. The results showed that patient voice agent-initial (VOS)
53 was judged as the most natural, followed by both orders in the agent voice. Patient voice
54 patient-initial (VSO) was judged as the least natural. Similar to Manuelli's results, in the
55
56
57
58
59
60

1
2
3 patient voice, ratings for the agent-initial condition (VOS) were statistically higher (more
4 grammatical) than the ratings for the patient-initial condition (VSO). In contrast, in the
5 agent voice, the ratings for the two orders were not statistically different from each other,
6 unlike the patient-initial (VOS) preference in Manuelli's study.
7
8

9 Using a picture description task, Sauppe, Norcliffe, Konopka, Van Valin, and
10 Levinson (2013) showed that adult speakers preferred to produce patient voice agent-
11 initial sentences (VOS, 62%), followed by agent voice patient-initial (VOS, 30%). The
12 speakers produced only few agent voice agent-initial (VSO, 5%), and patient voice
13 patient-initial (VSO, 2%). Tanaka (2016) used the same task but also manipulated
14 animacy and definiteness. The adult participants' preferences showed the same trend as
15 the preferences shown in Sauppe et al.'s study, but there was a smaller difference between
16 the frequency of agent voice patient-initial (VOS, 15%) and agent voice agent-initial
17 (VSO, 8%) productions.
18
19

20
21 The studies reported so far all indicate that patient voice agent-initial (VOS) is the
22 overall most preferred construction for adult speakers of Tagalog. The production data
23 further suggest that in the agent voice, subject final (i.e. agent final) are preferred
24 compared to agent-initial/subject-initial constructions. This is in line with the majority of
25 proposals on Tagalog's basic word order and supports the assumption that adults'
26 preferences are driven by an agent-first but also by a subject-last principle. The patient
27 voice agent-initial (VOS) structure obeys both of these principles and this could explain
28 the overall preference for these constructions. The results for the agent voice suggest a
29 slight dominance of the subject-last principle as the patient-initial and thus subject-last
30 order is preferred in this voice. This pattern could indicate that for Tagalog-speaking
31 adults grammatical principles of word order dominate principles that consider the
32 thematic roles of the arguments. However, such a conclusion is premature based on the
33 present data. In both production experiments, agent voice constructions were produced
34 less, so there were fewer data points to compare. Moreover, Sauppe et al. (2013) did not
35 control for the animacy of the themes in their stimuli. In Tanaka's study, there was no
36 agent voice production from stimuli with animate patients (N. Tanaka, personal
37 communication, February 15, 2017). As studies have shown that animacy has an effect on
38 word order choice and grammatical role assignment (Branigan, Pickering & Tanaka,
39 2007; Ferreira, 1994; Prat-Sala & Branigan, 2000), animacy should be controlled for in
40 an experiment in order to dissociate the subject-before-object from an agent-before-
41 patient preference.
42
43
44
45
46
47

48 Acquisition of Tagalog word order

49 Empirical studies on the acquisition of Tagalog word order are scarce. Bautista (1983)
50 used a picture description task to investigate children's (ages: 2;2–4;6) word order
51 preference. Her results showed a preference for the agent-initial order (88%) compared to
52 the patient-initial order (12%). In a longitudinal study of spontaneous speech samples of
53 six children (from age 1;2 to 4;11), Marzan (2013) reported that verb-agent-patient
54 constructions were one of the most used constructions in the data set. However, in neither
55 of these studies, voice was considered, so it cannot be determined based on these data
56 whether the agent-initial preference is dependent on voice.
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

More recently, Tanaka (2016) gave children (mean age: 5 years and 5 months) the same picture description task as in her production experiment with adults described above. According to her results, children-like adults-mostly produced patient voice agent-initial (VOS) constructions. However, in contrast to the adults, children preferred agent-initial utterances for the agent voice (VSO), as well. There were also a few patient voice patient-initial (VSO) constructions, but unlike the adults in her study, the children did not produce agent voice patient-initial (VOS) constructions.

The results from the reported studies on word order in Tagalog suggest differences between word order preferences of children and adults: Unlike adults who show an agent-initial preference only in the patient voice (VOS), children also seem to prefer agent-initial sentences in the agent voice (VSO). This suggests that children are less driven by the grammatical function of an argument but by the agent-first principle when choosing a word order in their production and thus would follow different principles in word order than adult speakers of their language. However, the data base on Tagalog-learning children is still too scarce to draw such a strong conclusion. First, due to the general preference of patient voice in children as well as adults, the number of utterances in agent voice was very limited in the previous production studies. Furthermore, Tanaka (2016) did not consider a potential effect on animacy on word order as both sentences with animacy contrast but also without animacy contrast of the arguments were included in her task but not analyzed separately. Since animacy has been shown to interact with children's word order preferences in other languages (Cannizzaro, 2012), we further investigated word order preferences in Tagalog using an experimental design that controlled for animacy and that provided conditions in which a higher number of agent voice productions could be elicited.

Current study

In this study, we wanted to investigate children's acquisition of word order patterns in a language wherein the first noun phrase position is not confounded with the subject position, and the agent is not closely linked to the subject. We examined Tagalog-learning children's word order preferences to determine whether an agent-first principle is stronger in guiding their word order preferences than a word order that is based on the grammatical function of the argument as data from adult Tagalog speakers suggest. We directly manipulated the voice-marking of the verbs in a sentence completion task to experimentally test and compare Tagalog adult's and children's word order preferences in agent and patient voice. Providing a voice-marked verb allowed us to investigate word order preferences in a highly controlled fashion and to elicit the same number of productions for the agent voice and the patient voice. The influence of animacy was also excluded by using only arguments without an animacy contrast. Furthermore, in order to see when children reach adult-like behaviour, we included adult participants and two groups of children that differed in age: 5-year olds as in the study by Tanaka (2016) and an older group of 7-year old children.

Experiment 1: Word order preferences of adults

In Experiment 1, we gave adult native speakers a sentence completion task to determine their word order preferences in both the agent and patient voice. Based on Kroeger's (1993) proposed agent-first and subject-last principles in ordering full noun phrases in Tagalog, the adults should show an agent-initial preference in the patient voice (VOS), but no such preference in the agent voice. However, if adults' word order preferences is more strongly guided by the subject-last than the agent first principle, we would expect more patient-initial than agent-initial orderings in the agent voice.

Method

Participants

Twenty native Tagalog speakers (mean age: 19 years, age range: 18–24 years, males: 10) participated in this study. They were all raised in Metro Manila, which was a selection criterion because there exist different Tagalog dialects in other Philippine provinces. No participant reported a history of language delay, or a psychiatric or neurologic disorder. Informed consent was obtained, participation was absolutely voluntary without any monetary compensation.

Materials

Sixteen causative 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 selected so that either of two animate entities could act as the agent or the patient. To keep animacy constant, we chose animals as doers and receivers of the actions. We assigned each verb to an animal pair from a pool of eight animals. Each verb was depicted in two pictures, such that the agent animal on the first picture was the patient animal on the second (see Figure 1 for an example of such a picture pair). The pictures were created by a professional artist. All of these images were digital, colored, and with a resolution of 1650 x 1276 pixels. We also counterbalanced the side on which the agent and patient appeared on the picture.

[insert Figure 1 about here]

The voice-marking of the verbs was manipulated, resulting in two conditions—agent voice and patient voice. For example, the verb *hila* 'pull' is *humihila* in the agent voice and *hinihila* in the patient voice condition. All verbs were inflected for the imperfective aspect, which is the easiest aspect for children to understand (Galang, 1982).

Verb-initial sentences containing the voice-inflected verbs were recorded in an audio recording booth using the Audacity(R) 2.1.0 program (Audacity Team, 2015) produced by a female native Tagalog speaker. The verbs were then cut from the wav files. Each sound file contained one inflected verb, had no silence, and was about 800ms long.

Procedure

Each item which consisted of one picture and its corresponding audio-recorded verb, was distributed into four different lists, following a Latin square design. In each list, each experimental condition appeared the same number of times (4), and each lexical verb was used only once, thus each list contained 16 experimental items. They were presented through DMDX Version 5 (Forster, K. I. & Forster, J. C., 2014), in a pseudo-randomized order, such that the same experimental condition was not presented more than three times consecutively.

The participants were tested in a quiet university room. The experimenter (first author) sat next to each participant, and presented the experiment on a 13-inch laptop which was about 50 centimeters away from the participant. The responses were recorded using a video or audio recorder.

First, the experimenter presented single pictures of the animals that would appear in the main experiment, as well as the actions mentioned in the stimuli sentences (the pictures showed the actions between two boys instead of between two different animals like in the main experiment). The pictures were presented four at a time, and the participants were asked to point to the picture of the concept that was labeled by the experimenter. This task was administered to ensure consistency with Experiment 2, in which children were tested. Next, the sentence completion task was conducted. The participants were informed that they would first see a picture, and then hear a word through the head phones. Their task was to complete the sentence which starts with this voice-marked verb, in order to describe the scene depicted in the picture. They were told not to repeat the verb, and only mention the arguments. Each picture was presented in full screen for 2500ms before the audio-recorded verb was played. The picture remained on the screen and the audio was replayed after every 10 seconds as long as no response was provided.

Four practice items were presented before the actual experiment. Feedback was given, but was limited to reminders that the given word (the verb) should be in the beginning of the sentence, and that the event in the picture should be completely described. During the actual experiment, no feedback was given. The participants were offered a chance to have a break halfway through the experiment.

Data analysis

The video and audio recordings of the testing sessions were transcribed by a native Tagalog speaker. The independent variable was voice marking (agent voice, patient voice), and the depicted agent's position in the sentence was treated as dependent variable. The statistical analysis software R version 3.2.5 (R Core Team, 2016) was used for computations. Chance performance was analyzed using logistic mixed models, specifically the R function `glmer` (family = binomial, optimizer = bobyqa) of the `lme4` package version 1.1-12 (Bates, Maechler, Bolker, & Walker, 2016). The model included the fixed effect of voice marking, and the random effect of voice marking by subject and by item.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Only sentence completions that involved the presented verbs and two arguments were included into the data analysis. Items including the verb *pasan* 'give a piggyback ride' were removed from all subsequent analysis due to an unexpected thematic role assignment by several participants. For example, for a picture of a chicken giving a mouse a piggyback ride, we expected that in the agent voice the *ang* marker would be used for the chicken as it is the agent of the action. However, 6 out of the 20 adult participants used *ang* for the mouse, and the preposition *sa* instead for the chicken which turned it into a locative, which means that the mouse is the agent, doing the action of riding the chicken. We excluded cases of other incorrect verb interpretations, i.e., use of *karga* 'carry' to mean *talon* 'jump unto' (3%), and single-argument productions (3%). In addition, there were instances of a mismatch between the noun markings and the action in the picture (1%): reversals of the markers *ang* and *ng*, and use of the *ang* marker for both arguments. However, these were still included in the chance-level testing. The results do not differ when these instances are excluded.

Results

We analyzed whether adults' production of agent-initial constructions exceeded chance-level, i.e., 50% as the agent could occur only before or after the patient. The logistic mixed model showed that the amount of agent-initial constructions was not different from chance in the agent voice (Estimate = 0.003, SE = 0.59, $z = 0.005$, $p > .99$), but above chance in the patient voice condition (Estimate = 6.31, SE = 2.37, $z = 2.67$, $p = .008$) (see Figure 2).

[insert Figure 2 about here]

Discussion

Adults showed an agent-initial preference in the patient voice (VOS), but no preference for one of the orders in the agent voice, showing that voice affects word order preferences in Tagalog speakers. The agent-initial preference in the patient voice is in line with findings from previous studies on Tagalog, which utilized other methods, such as grammaticality judgement and free picture description tasks (Manueli, 2010; Hsieh, 2016; Sauppe et al, 2013; Tanaka, 2016).

The lack of a clear preference in the agent voice, which is in line with Hsieh's (2016) findings, provides empirical support for Kroeger's claim that the order of full noun phrases in Tagalog is guided not only by an agent-first principle but also by a subject-last principle. In the agent voice these two principle are in competition with each other as an agent-initial construction (VSO) satisfies only the agent-first principle, but violates the subject-last principle while a patient-initial construction (VOS) satisfies only the subject-last but not the agent-first principle. This competition could explain why no preferred word order could be found in this condition. In contrast, in the patient voice, both principles are satisfied in an agent-initial construction (VOS), but violated in a patient-

1
2
3 initial construction (VSO) which may lead to a high and homogeneous preference of the
4 agent-initial and simultaneously subject-last order in this condition.
5
6

7 However, the results do not fully conform to Sauppe et al.'s (2013) and Tanaka's
8 (2016) findings which seemed to show that speakers put more weight on the subject-last
9 principle compared to the agent-first principle. In contrast, in the current experiment, no
10 principle seems to outweigh the other as no statistically significant preference for one or
11 the other order was found in our results for the agent voice. As already mentioned, our
12 study did not elicit productions with animacy contrasts between the arguments, so
13 differences in the results between our study and those of Sauppe et al. and Tanaka can be
14 due to the fact that they did not control for this factor. Since animacy has been shown to
15 affect both word order and grammatical functions (Branigan, Pickering & Tanaka, 2007;
16 Ferreira, 1994; Prat-Sala & Branigan, 2000), the use of animate patients in the current
17 study and mostly inanimate patients in the previous studies makes it difficult for the
18 results to be compared with each other.
19
20
21
22
23

24 Experiment 2: Word order preferences of children

25
26
27 In Experiment 2, five- and seven-year-old Tagalog children were tested with the same
28 experimental design and the same material as the Tagalog speaking adults in Experiment
29 1. We wanted to know whether Tagalog-speaking children have a general agent-initial
30 bias, similar to children learning other languages; or if they show this preference only in
31 the patient voice, similar to the adults in Experiment 1.
32
33

34 Method

35 Participants

36
37
38 In total, 65 typically-developing children from Metro Manila, Philippines, participated in
39 the study. There were 34 five-year-old and 31 seven-year-old participants. All of the
40 children had Tagalog as their dominant language. The five-year-old children (M : 5;9,
41 range: 5;4–5;11, males: 11) were Kindergarten students from two public elementary
42 schools, while the seven-year-olds (M : 7;8, range: 7;3–7;11, males: 20) were Grade 2
43 students from the same schools. Informed consent was obtained from the children's
44 parents. No history of language delay was reported for any of the children.
45
46
47
48
49

50 Materials

51 The materials were identical to those used in Experiment 1.
52

53 Procedure

54 The procedure was similar to that of Experiment 1, with a few additional instructions for
55 the children. Whenever a participant had made a mistake during the pre-experiment phase
56 where they had to identify the animals and actions used in the experiment, they were
57
58
59
60

1
2
3 reminded to look at all of the pictures again, and to listen more carefully. They were then
4 asked to identify all the items on the screen again. The practice session for the actual
5 experiment was started only when the participant had succeeded in identifying all the
6 animals and actions.
7
8

9 There were four practice items. During this phase, feedback was given, but was
10 limited to reminders that the given word (the verb) was the beginning of the sentence,
11 that the word should not be changed, and that the event in the picture should be
12 completely described. No corrections were given when the participants used
13 morphological markers on the nouns which did not match the event depicted in the
14 picture. During the actual experiment, no feedback was given except when the
15 participants changed the inflection of the verb (most of the children repeated the verb to
16 start their sentence). These incorrectly repeated items were presented again, by waiting
17 for the 10-second programmed time for the verb prompt to be replayed. In addition, to
18 motivate the children to finish the task, the experiment was presented as a game, in which
19 they had to help a boy reach the finish line in a race.
20
21
22
23

24 Data analysis

25 Data analysis followed Experiment 1, with the addition of age group (five-year-olds [5-
26 yo], seven-year-olds [7-yo]) as an independent variable. The correctness of the
27 morphological markers on the noun phrases, and the specific errors made (morphological
28 marker reversals, or the use of *at* 'and' to conjoin the two noun phrases, and of *ang* or *ng*
29 for both noun phrases) were also noted. Self-corrections were considered as correct only
30 when the children produced single-argument constructions on the first try (twice in the
31 five-year-olds, four times in the seven-year-old group). The trials in which the
32 participants changed the voice-marking on the verb were not included. The item *pasan*
33 'give a piggyback ride' was also excluded from subsequent analyses because of the same
34 reason as in Experiment 1.
35
36
37
38

39 Aside from the *pasan* items, 9% of the remaining data points were excluded
40 because of the following reasons: single-argument productions (4% from the 5-yo, 2%
41 from the 7-yo), incorrect interpretation of the verb (0.9% from the 5-yo, 0.4% from the 7-
42 yo), conjoined noun phrase productions (0.5% from the 5-yo, 0.7% from the 7-yo), not
43 following the instructions (0.5% from the 5-yo, 0.1% from the 7-yo), and skipped trials
44 (0.1% from the 5-yo, and 0.2% from the 7-yo).
45
46

47 In addition, in 14% of the data, there was a mismatch between the noun markers
48 and the action in the picture (8% from the 5-yo, 6% from the 7-yo). For example, given
49 an agent voice inflected verb, a participant marked pig with *ang*, when the agent in the
50 picture was the cow, and thus, should have had the *ang* marker. The breakdown of the
51 errors is discussed in a separate section. However, these were still included in the model,
52 as we were mostly interested in the order of mention of the agent of the action depicted in
53 the picture (see Appendix for the results of the analyses when nouns with mismatched
54 markers were excluded). The number of analyzed data points per condition is shown in
55 Table 1.
56
57
58
59
60

Table 1. Number of analyzed data points per condition in Experiment 2

	Five-year-olds	Seven-year-olds
Agent voice	217	214
Patient voice	237	222

Results

We analyzed whether 5-yo and 7-yo produced more agent-initial constructions in each voice condition compared to chance (see Figure 3). The logistic mixed model showed that both age groups more often used an agent-initial construction than expected by chance in both voice-marking conditions in both age groups (for details, see Table 2).

[insert Figure 3 about here]

Table 2. Results of chance-level testing using a logistic mixed model on children's word order preference in Experiment 2

Predictor	Estimate	Standard error	z value	p value
5-yo agent voice	2.37	0.35	6.72	<.001
5-yo patient voice	1.60	0.40	4.03	<.001
7-yo agent voice	2.72	0.39	6.90	<.001
7-yo patient voice	2.43	0.46	5.25	<.001

Note. 'yo' = year-olds.

As noun marking errors provide insight on the children's mastery of the Tagalog voice-marking system and their preferred voice marking on the verb, we also analyzed the accuracy in marking the nouns with *ang* and *ng* (see Figure 4). Accurate means that both nouns were marked correctly in relation to the voice-marking on the verb, such that the sentence interpretation matches the action depicted on the picture. We fitted a mixed-effects logistic model to determine the effects of voice, age, and their interaction, on the accuracy in marking the nouns with *ang* and *ng*. The results show a main effect of voice (see Table 3 for a summary), such that the children were more accurate in marking the nouns given a patient voice-inflected verb compared to an agent voice-inflected verb.

[insert Figure 4 about here]

Table 3. Summary of the fixed effects of voice, age, and their interaction on children's noun-marking accuracy in Experiment 2 (N = 890; log-likelihood = -327.5)

Predictor	Estimate	Standard error	z value	p value
Intercept	2.66	0.31	8.50	<.001
Voice	1.16	0.54	2.17	.03
Age	0.36	0.50	0.72	.47
Voice * Age	-0.31	0.38	-0.80	.42

The 5-yo produced 17% of the nouns with a wrong marking while the 7-yo made only 14% of such errors. In the 5-yo, 66% of the errors occurred in the agent voice, while this was the case for 52% of the errors in the 7-yo. For both 5-yo and 7-yo, more than half of the errors were reversals of *ang* and *ng* (see Figure 5 for a breakdown of the errors per age and voice-marking). The remaining errors consisted of the use of *ang* for both arguments or *ng* for both arguments, and of dropping an article in one of the arguments.

[insert Figure 5 about here]

We also further analyzed the reversals of *ang* and *ng*. We fitted a mixed-effects logistic model to determine the effects of age, voice, and the interaction of the two, on the number of reversal errors. The results showed no significant main effects of age, voice, nor the interaction of the two (see Table 4 for a summary of the results).

Table 4. Summary of the fixed effects of voice, age, and their interaction on children's noun-marking reversal errors in Experiment 2 (N = 890; log-likelihood = -220)

Predictor	Estimate	Standard error	z value	p value
Intercept	-4.15	0.69	-6.06	<.001
Voice	-2.35	1.36	-1.73	.08
Age	-0.18	0.62	-0.29	.77
Voice * Age	0.26	0.65	0.39	.69

Out of the 34 5-yo participants, 30 repeated the verb before completing the sentence, while 25 out of the 31 7-yo did so. The 5-yo incorrectly repeated the verb in 7% of the total number of experimental trials, and the 7-yo in 8% of the trials. For the 5-yo, 79% of these incorrect repetitions involved a change of an agent voice-inflected verb to a patient voice-inflected verb, 14% of changes from patient voice to agent voice, and 7% of instances of use of other verbs aside from the one provided which was *sagip* 'rescue.' For

1
2
3 the 7-yo, 90% of the incorrect repetitions involved a change from agent voice to patient
4 voice, 3% change from patient voice to agent voice, and 7% of use of other patient voice
5 inflected verb for *sagip* 'rescue.'
6
7
8
9

10 Discussion

11
12 Five- and seven-year-old children showed a preference for agent-initial constructions in
13 both the agent voice (VSO) and the patient voice (VOS). The same preference was found
14 for the five-year-old Tagalog-speaking children in Tanaka's (2016) free picture
15 description experiment. Moreover, these results are in line with findings of an early
16 agent-initial preference in other languages (Angiolillo & Goldin-Meadow, 1982; Brown,
17 1973 for English; Cannizzaro, 2012 for Dutch and English; Lee, 2010 for Mandarin;
18 Slobin & Bever, 1982 for Italian, Serbo-Croatian, and Turkish; Tanaka & Shirai, 2012 for
19 Japanese). The current results also support Jackendoff and Wittenberg's (2014) proposal
20 that children prefer a word order that places the agent before the patient.
21
22
23

24
25 More importantly, these findings support the claim that this word order preference
26 results from a direct mapping of thematic roles and linear argument order, without regard
27 for grammatical categories like subject and object, and their preferred positions. In
28 Tagalog, an agent-before-patient order is the same as a subject-before-object order only
29 in the agent voice. However, in the current experiment, an agent-initial preference was
30 also found in the patient voice, which has an object-before-subject order. As the first
31 noun phrase position in Tagalog is not confounded with the subject position, and the
32 agent is not a preferred subject, the results show that children's word order preference is
33 semantically-driven, and not merely brought about by the preferred order of grammatical
34 functions (i.e., subject-before-object). The current findings support claims that the agent-
35 initial preference is due to more general and not exclusively linguistic reasons such as the
36 agent's higher ranking in the thematic hierarchy (Siewierska, 1993), and thematical
37 independence (Primus, 2006).
38
39

40
41 The results further revealed that children used the appropriate noun markings in
42 both the agent and the patient voice in more than 75% of their productions. However,
43 they still made noun marking errors, such as the use of *ang* instead of *ng* or vice versa, or
44 not using a noun marker at all, which may indicate that they have not yet fully mastered
45 the voice-marking system of Tagalog. However, we attribute these errors to the design of
46 the experiment. Children might have anticipated a particular voice-marked verb upon
47 seeing the action in the picture, and could not overwrite this when they heard another
48 voice-marking instead. The errors in noun-marking and the exchange of the voice-marker
49 on the verb provide insight on this anticipation or voice preference. For both groups of
50 children, there were more noun-marking errors in the agent voice than in the patient
51 voice, and the incorrect repetitions of the voice-marking of the verbs were mostly
52 changes of the agent voice inflection to the patient voice compared to the reverse. These
53 findings implicate a general patient voice preference, which has also been previously
54 found in a less restricted picture description task for children as well as adults (Tanaka,
55 2015; 2016).
56
57
58
59
60

General Discussion

This study investigated the word order preferences of Tagalog-speaking adults and children. The results of Experiment 1 showed that adults prefer agent-initial constructions in the patient voice (VOS), but they had no word order preference in the agent voice, supporting previous research on word order preferences in Tagalog (Hsieh, 2016). The results are also in line with Kroeger's (1993b) proposed principles guiding the ordering of non-pronominal arguments in Tagalog: agent-first and subject-last.

In Experiment 2, the findings revealed that five- and seven-year-old children preferred agent-initial constructions not only in the patient voice (VOS), but also in the agent voice (VSO)—a pattern different from the one shown by the adults. The children's data are consistent with Tanaka's (2016) findings, supporting the claim that children exhibit this universal tendency of an agent-initial preference early on (Jackendoff & Wittenberg, 2014). The results do not show a subject-before-object preference, but an agent-initial preference by the children instead. Primus (2006) proposed that an agent-initial preference is due to a universal principle that the thematically independent role (agent) tends to precede and/or c-command the role that is thematically dependent (patient). Our data from Tagalog-speaking children support this assumption and show that this preference is quite stable in children even if their language does not provide unique support for this ordering.

The results also implicate that adult's word order preferences are affected by the voice-marking on the verb. Adults show a preference for agent-initial orderings only in the patient voice, but not in the agent voice. This suggests that Tagalog has a canonical word order only in the patient voice, which can be characterized as agent-initial and subject-final. If children first acquire the language's canonical forms, (Slobin, 1982; Slobin & Bever, 1982), Tagalog-speaking children would first acquire the patient voice agent-initial order. The results of the current study support this claim. As regards the children's agent-initial preference in the agent voice, it can be the case that they have derived this pattern from the canonical order for the patient voice. After all, the patient voice is also more frequently used compared to the agent voice in constructions with two arguments. When we analyzed a child-directed speech sample taken by Marzan (2013) from daily family interactions of one Tagalog-speaking child (from the age of 3:0 to 4:11), 83% of constructions with voice-inflected verbs and two arguments were in the patient voice, and only 17% were in the agent voice. The patient voice constructions were also 90% agent-initial. The results of the current study imply that children overgeneralize this canonical order for the patient voice to the agent voice.

It can be argued that children have the same word order preference for the two voices, only because they could not distinguish and make use of the verb markers for the agent and for the patient voice. However, children's correct use of the noun markers in both voices in the current experiment actually shows that they know that the agent voice infix on the verb marks the *ang*-phrase as the agent of the action while the patient voice infix on the verbs marks the *ang*-phrase as the patient. Their preference of an agent-initial

1
2
3 construction for both voices also shows that they are aware of the differences between the
4 two voices, since agent-initial is VSO in the agent voice, but VOS in the patient voice.
5
6

7 Children's agent-initial preference for both voices shows that they rely more
8 strongly on the agent-first principle, than on the subject-last principle for ordering non-
9 pronominal arguments in Tagalog. It can be that children are aware of the two principles,
10 but give priority to the agent-first principle compared to the subject-last principle.
11 Another possibility is that even seven-year-old Tagalog-speaking children follow only the
12 more universal agent-first principle, but have not yet acquired the more language-specific
13 subject-last principle. This suggests that the acquisition of some language-specific
14 features may fairly go beyond the age of seven years.
15
16

17 The lower priority or the late acquisition of the subject-last principle may also be
18 due to the low frequency of utterances with two full noun phrases which is typical of
19 spontaneous speech in general (Du Bois, 1987). If most utterances contain a pronoun, the
20 subject-last principle may not be well-attested in the language input that children receive,
21 given the more constrained order of pronouns in Tagalog. Given verb-initial sentences,
22 pronouns are expected to appear immediately after a verb irrespective of their
23 grammatical function (Billings, 2005). In fact, in the same child-directed speech sample
24 from Marzan (2013), we found that only 3% of the utterances with a voice-inflected
25 transitive verb had two non-pronominal arguments. Among these utterances, only one
26 had a subject-last order. The rest of the utterances with two arguments contained
27 pronouns. In 97% of sentences with one pronoun and one full noun phrase, the pronoun
28 referred to the agent. Therefore, it can be that Tagalog-speaking children hear very few
29 non-pronominal constructions with a subject-last order.
30
31
32

33 Children may prefer agents in the initial position, even in the agent voice, because
34 they overgeneralize the stricter order of pronouns in their input. As shown by Matthews,
35 Lieven, Theakston, and Tomasello (2005), children may learn word order from
36 distributionally regular items like pronouns. Since agents are usually given information in
37 discourse (DuBois, 2003), and pronouns are used to represent given information (Weber
38 & Mueller, 2004), agents tend to be coded as pronouns (Bowerman, 1978). There is even
39 evidence that pronouns tend to be used for agents than for patients (Angiolillo & Goldin-
40 Meadow, 1982). Since pronouns occupy the position immediately after the verb in
41 canonical Tagalog sentences (Billings, 2005), agent-initial constructions are probably
42 more frequent in the input if pronouns are typically referring to agents. In the same child-
43 directed speech sample from Marzan (2013) as described above, 86% of the two-
44 argument utterances which contained at least one pronominal argument had an agent-
45 initial order. In addition, all of the agent voice utterances with pronouns were actually
46 agent-initial. We can say that children may have overgeneralized the dominant agent-
47 initial order of sentences with pronouns in the input to sentences with non-pronominal
48 arguments like in the current experiment.
49
50
51
52

53 Overall, the results show that children were aware of the flexibility of word order
54 in Tagalog, since they preferred VSO in the agent voice but VOS in the patient voice.
55 However, at age seven, they still did not show adult-like distribution of productions in the
56 agent voice. Concerning the children's ability to use voice-marking per se, there were
57
58
59
60

more instances of incorrect repetitions of the verbs from the agent voice to the patient voice, compared to the opposite direction. Moreover, both groups of children made more errors in marking the nouns in the agent voice compared to the patient voice. These results implicate a patient voice preference which is in line with findings from adult corpus data (Cooreman, Fox, & Givon, 1984) and previous production experiments (Sauppe et al., 2013; Tanaka, 2015, 2016). These show that with respect to the distribution of the agent and patient voice, even five-year-old children are more adult-like than with respect to word order properties. It can be that the complex interplay of verbal and nominal markings and word order in Tagalog makes the system more difficult for children to acquire.

It is then of interest to test older children to see when they start showing adult-like distributions. In addition, cross-linguistic comprehension studies show that children start correctly interpreting non-canonical word orders (e.g., use of morphological markers instead of a first-noun-phrase-as-agent strategy) as early as around two years of age in Turkish (Slobin & Bever, 1982), and between the age of five and seven in German (Dittmar et al., 2008). Given the strong preference for agent-initial productions in Tagalog even at the age of seven, it is interesting to see whether this preference is also found in comprehension.

In conclusion, even at the age of seven, Tagalog-speaking children are still tuning into the word order preferences of their language. A lot remains to be explored in the acquisition of word order, and investigating this in understudied languages can broaden our understanding of this phenomenon.

Appendix

Analysis on productions with correctly-marked nouns only

Analysis of the data without the 17% instances of a mismatch between the noun markings and the action in the picture in the five-year-old group, and 14% in the seven-year-old group shows similar results to that of the analysis including all of the productions. The fitted logistic mixed model showed above chance level agent-initial productions in both voice conditions for both age groups (see Table 5), just like the main model.

Table 5. Results of chance-level testing using a logistic mixed model on children's word order preference excluding incorrectly-marked nouns in Experiment 2

Predictor	Estimate	Standard error	z value	p value
5-yo agent voice	2.72	0.45	6.11	<.001
5-yo patient voice	1.94	0.54	3.58	<.001
7-yo agent voice	3.16	0.51	6.20	<.001

7-yo patient voice	2.60	0.63	4.14	<.001
--------------------	------	------	------	-------

Note. 'yo' = year-olds.

Notes

1. Voice-marking and mood are conflated in Tagalog verbs. In this work, the voice-markings used also signal realis mood. See Himmelmann (2005b) for a longer discussion on voice-marking and mood in Tagalog.
2. Billings (2005) claim that the agent-first principle actually applies only to given agents. The effect of givenness on Tagalog speakers' word order preference is beyond the scope of this study.

References

- Abbot-Smith, K. & Serratrice, L. (2015). Word order, referential expression, and case cues to the acquisition of transitive sentences in Italian. *Journal of Child Language*, 42, 1-31.
- Aldridge, E. (2002). Nominalization and WH-movement in Seediq and Tagalog. *Language and Linguistics*, 3(2), 393-426.
- Audacity Team (2015). Audacity(R): free audio editor and recorder (Version 2.1.0) [computer program]. Retrieved from <<http://audacity.sourceforge.net/>>.
- Angiolillo, C. J. & Goldin-Meadow, S. (1982). Experimental evidence for agent-patient categories in child language. *Journal of Child Language*, 9(3), 627-643.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. Retrieved from <<http://CRAN.R-project.org/package=lme4>>.
- Bautista, M. L.S. (1983). Word order in Tagalog child language. *Philippine Journal of Linguistics*, 14(1), 25-44.
- Billings, L. A. (2005). Ordering clitics and postverbal R-expressions in Tagalog: a unified analysis? In A. Carnie, H. Harley & S. A. Dooley (eds.), *Verb first: on the syntax of verb initial languages*, 303-339. Amsterdam: John Benjamins.
- Bowerman, M. (1978). Words and sentences: uniformity, individual variation, and shifts over time in patterns of acquisition. In F. Minifie & L. Lloyd (eds.), *Communicative and cognitive abilities - early behavioral assessment*. Baltimore: University Park Press
- Bronckart, J. P., Gennari, M. & De Weck, G. (1981). The comprehension of simple sentences: The representative perspective and the communicative perspective. *International Journal of Psycholinguistics*, 8, 5-29.
- Brown, R. (1973). *A first language: the early stages*. London: George Allen & Unwin.
- Cannizzaro, C. L. (2012). *Early word order and animacy*. Groningen: Groningen Dissertations in Linguistics 104.

- 1
2
3 Chan, A., Lieven, E. & Tomasello, M. (2009). Children's understanding of the agent-
4 patient relations in the transitive construction: cross-linguistic comparisons
5 between Cantonese, German, and English. *Cognitive Linguistics*, 20(2), 267-300.
- 6 Cohn, N. & Paczynski, M. (2013). Prediction, events, and the advantage of agents: The
7 processing of semantic roles in visual narrative. *Cognitive Psychology*, 67, 73-97.
- 8 Cooreman, M., Fox, B. & Givón, T. (1984). The discourse definition of ergativity.
9 *Studies in Language*, 8(1), 1-34.
- 10 Dittmar, M., Abbot-Smith, K., Lieven, E. & Tomasello, M. (2008). Young German
11 children's early syntactic competence: a preferential looking
12 study. *Developmental Science*, 11, 575-582.
- 13 Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath
14 (eds.), *The world atlas of language structures online*. Leipzig: Max Planck
15 Institute for Evolutionary Anthropology. Retrieved from
16 <<http://wals.info/chapter/81>>.
- 17 Du Bois, J. W. (1987). The discourse basis of ergativity. *Language* 64, 805-855.
- 18 Du Bois, J. W. (2003). Argument structure: Grammar in use. In J. W. Du Bois, L. E.
19 Kumpf, & W. J. Ashby (eds.), *Preferred argument structure: Grammar as*
20 *architecture for function*. Amsterdam/Philadelphia: Benjamins, 11-60.
- 21 Fillmore, J. (1968). The case for case. In E. Bach & R. Harms (eds.), *Universals in*
22 *linguistic theory*. New York: Rinehart & Winston.
- 23 Forster, K. I., & Forster, J. C. (2014). DMDX (Version 5) [computer program]. Tucson:
24 University of Arizona. Retrieved from
25 <<http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>>.
- 26 Galang, R. (1982). Acquisition of Tagalog verbal morphology: linguistic and cognitive
27 factors. *Philippine Journal of Linguistics*, 13, 1-15.
- 28 Greenberg, J. H. (1963). Some universals of grammar with particular reference to the
29 order of meaningful elements. In J. H. Greenber (ed.), *Universals of language*, 73-
30 113. Cambridge, MA: MIT Press.
- 31 Guilfoyle, E. Hung, H. & Travis, L. (1992). Spec of IP and spec of VP: two subjects in
32 Austronesian languages. *Natural Language and Linguistic Theory*, 10, 375-414.
- 33 Hakuta, K. (1982). Interaction between particles and word order in the comprehension
34 and production of simple sentences in Japanese children. *Developmental*
35 *Psychology*, 18(1), 62-76.
- 36 Himmelmann, N. P. (2004). Tagalog. In G. Booij, C. Lehmann, J. Mugdan & S.
37 Skopeteas (eds.), *Morphology: an international handbook on inflection and word*
38 *formation*, 1473-1490. Berlin: de Gruyter.
- 39 Himmelmann, N. P. (2005a). The Austronesian languages of Asia and Madagascar:
40 typological characteristics. In K. A. Adelaar & N. P. Himmelmann (eds.), *The*
41 *Austronesian languages of Asia and Madagascar*, 110-181. London: Routledge.
- 42 Himmelmann, N. P. (2005b). Tagalog. In K. A. Adelaar & N. P. Himmelmann (eds.),
43 *The Austronesian Languages of Asia and Madagascar*, 350-376. London:
44 Routledge.
- 45 Hsieh, H. (2016). Prosodic indicators of phrase structure in Tagalog transitive sentences.
46 In H. Nomoto, T. Miyauchi & A. Shiohara (eds.), *AFLA 23: The proceedings of*
47 *the 23rd meeting of the Austronesian Formal Linguistics Association*. Canberra:
48 A.C.T. Asia-Pacific Linguistics.
- 49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Jackendoff, R. & Wittenberg, E. (2014). What can you say without syntax: a hierarchy of
4 grammatical complexity. In F. Newmeyer & L. Preston (eds.), *Measuring*
5 *linguistic complexity*, 65– 82. Oxford: Oxford University Press.
- 6
7 Johnson, D. E. (1977). On relational constraints on grammars. In P. Cole & J. M. Sadock
8 (eds.), *Syntax and Semantics 8: Grammatical relations*, 151- 178. New York:
9 Academic Press.
- 10
11 Kemmerer, D. (2012). The cross-linguistic prevalence of SOV and SVO word orders
12 reflects the sequential and hierarchical representation of action in Broca's area.
13 *Language and Linguistics Compass*, 6(1), 50–66.
- 14
15 Kroeger, P. (1993a). Another look at subjecthood in Tagalog. *Philippine Journal of*
16 *Linguistics*, 24(2), 1-16.
- 17
18 Kroeger, P. (1993b). *Phrase structure and grammatical relations in Tagalog*. Stanford:
19 CSLI Publications.
- 20
21 Latrouite, A. (2001). Linking in Tagalog: argument encoding determined by the semantic
22 properties of arguments. In G. Booij & J. van Marle (eds.), *Yearbook of*
23 *morphology 2000*, 123-154. Kluwer Academic Publishers.
- 24
25 Lee, T. H. T. (2010). The acquisition of word order in a topic-prominent language:
26 Corpus findings and experimental investigation. In R. Otaguro, K. Ishikawa, H.
27 Umemoto, K. Yoshimoto, & Y. Harada (eds.), *Proceedings of the 24th Pacific*
28 *Asia Conference on Language, Information and Computation*, 15. Tohoku:
29 Institute for Digital Enhancement of Cognitive Processing, Waseda University.
- 30
31 Lewis, M. P., Simons, G.F. & Fennig, C. D. (2016). *Ethnologue: languages of the world*,
32 Nineteenth edition. Dallas, Texas: SIL International. Retrieved from
33 <<http://www.ethnologue.com>>.
- 34
35 Manuelli, M. K. S. (2010). A minimalist account of scrambling and word order in
36 Tagalog: Testing its grammaticality. *Philippine Social Science Review (PSSR)*,
37 62(2), 161-195.
- 38
39 Marzan, J.C. (2013). *Spoken language patterns of selected Filipino toddlers and*
40 *preschool children*. Unpublished doctoral dissertation, University of the
41 Philippines Diliman.
- 42
43 Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency
44 in the acquisition of English word order. *Cognitive Development*, 20(1), 121-136.
- 45
46 McFarland, C. D. (1989). A frequency count of Filipino. *Linguistic Society of the*
47 *Philippines*.
- 48
49 Naylor, P.B. (1975). Topic, focus, and emphasis in the Tagalog verbal clause. *Oceanic*
50 *Linguistics*, 14(1), 12-79.
- 51
52 Primus, B. (2003). Division of Labour: The role-semantic function of basic order and
53 case. In D. Willems, B. Defrancq, T. Coleman & Noël, D. (eds.), *Contrastive*
54 *analysis in language*, 89-133. New York: Palgrave Macmillan.
- 55
56 Primus, B. (2006). Hierarchy mismatches and the dimensions of role semantics. In I.
57 Bornkessel, M. Schlesewsky, B. Comrie & A. Friederici (eds.), *Semantic role*
58 *universals and argument linking: theoretical, typological and psycholinguistic*
59 *perspectives*, 53-88. Berlin: Walter de Gruyter.
- 60
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/>>.

- 1
2
3 Rackowski, A. (2002). The structure of Tagalog: specificity, voice, and the distribution of
4 arguments. Massachusetts: Massachusetts Institute of Technology.
5
6 Richards, N. (1981). Another look at Tagalog subjects. In I. Paul, V. Philips & L. Travis
7 (eds.), *Formal issues in Austronesian linguistics*. Dordrecht: Kluwer.
8
9 Riesberg, S. (2014). Symmetrical voice and linking in western Austronesian languages
10 (*Pacific linguistics* 646). Boston/Berlin: De Gruyter Mouton.
11
12 Riesberg S. & Primus B. (2015). Agent prominence in symmetrical voice
13 languages. *STUF - Language Typology and Universals*, 68, 551–564.
14
15 Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin Jr., R. D. & Levinson, S. C.
16 (2013). Dependencies first: eye tracking evidence from sentence production in
17 Tagalog. In M. Knauff, M. Pauen, N. Sebanz & I. Wachsmuth (eds.), *Proceedings*
18 *of the 35th annual meeting of the Cognitive Science Society (CogSci 2013)*,
19 1265-1270. Austin, TX: Cognitive Science Society.
20
21 Sauermann, A. (2016). Impact of the type of referring expression on the acquisition of
22 word order variation. Potsdam: Universitätsverlag Potsdam.
23
24 Schachter, P. (2015). Tagalog. In T. Kiss & A. Alexiadou (eds.), *Syntax - theory and*
25 *analysis: an international handbook (Handbooks of Linguistics and*
26 *Communication Science Vol. 42/3)*, 1658-1676. Berlin: Mouton de Gruyter.
27
28 Schachter, P. & Otnes, F.T. (1972). *Tagalog Reference Grammar*. Berkeley: University
29 of California Press.
30
31 Segalowitz, N. S. & Galang, R. S. (1978). Agent–patient word-order preference in the
32 acquisition of Tagalog. *Journal of Child Language*, 5, 47-64.
33
34 Siewierska, A. (1993). On the interplay of factors in the determination of word order. In
35 J. Jacobs et al. (eds.), *Syntax: an international handbook of contemporary*
36 *research*, 826-846. Berlin: Walter de Gruyter.
37
38 Slobin, D. I. (1982). Universal and particular in the acquisition of language. In L. R.
39 Gleitman & E. Wanner (eds.), *Language acquisition: state of the art*, 128-170.
40 Cambridge: Cambridge University Press.
41
42 Slobin, D. I. & Bever, T. G. (1982). Children use canonical sentence schemas: A
43 crosslinguistic study of word order and inflections. *Cognition*, 12, 229-65.
44
45 Tanaka, N. (2016). An asymmetry in the acquisition of Tagalog relative
46 clauses. Unpublished doctoral dissertation, University of Hawaii at Manoa.
47
48 Tanaka, N., O’Grady, W., Deen, K., Kim, C. E., Hattori, R., Bondoc, I. P. M., & Soriano,
49 J. U. (2015). Split focus preferences in Tagalog: evidence from child language. In
50 A. Camp, Y. Otsuka, C. Stabile & N. Tanaka (eds.), *AFLA 21: the proceedings of*
51 *the 21st meeting of the Austronesian Formal Linguistics Association*, 279–288.
52 Canberra: Asia-Pacific Linguistics.
53
54 Tanaka, N. & Shirai, Y. (2014). L1 acquisition of Japanese transitive verbs: How do
55 children acquire grammar in the absence of clear evidence? In S. Nam, H. Ko & J.
56 Jun (eds.), *Japanese/Korean linguistics (Vol. 21)*, 281-295. Stanford, CA: CSLI.
57
58 Van der Lely, H. K. J. (1994). Canonical linking rules: forward versus reverse linking in
59 normally developing and specifically language-impaired children. *Cognition*, 51,
60 29–72.
61
62 Weber, A., & Mueller, K. (2004). Word order variation in German main clauses: A
63 corpus analysis. In *COLING ‘04 Proceedings of the 20th International*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Conference on Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics.

For Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

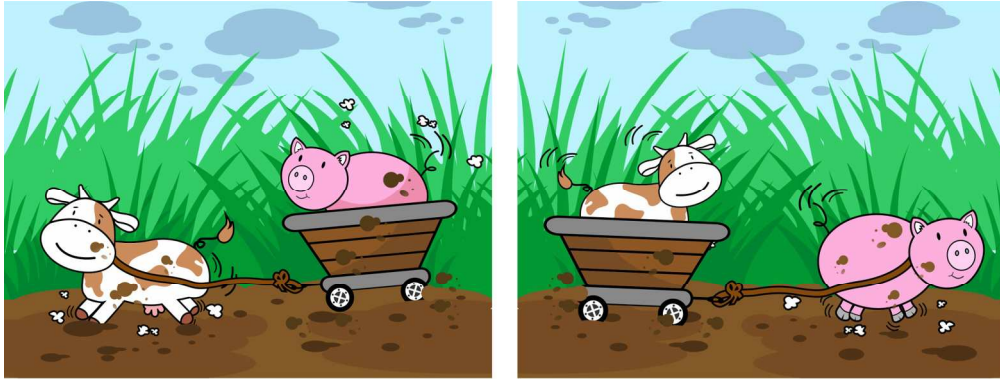


Figure 1. Picture pair for the verb—*hila* 'pull.'

694x261mm (72 x 72 DPI)

Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

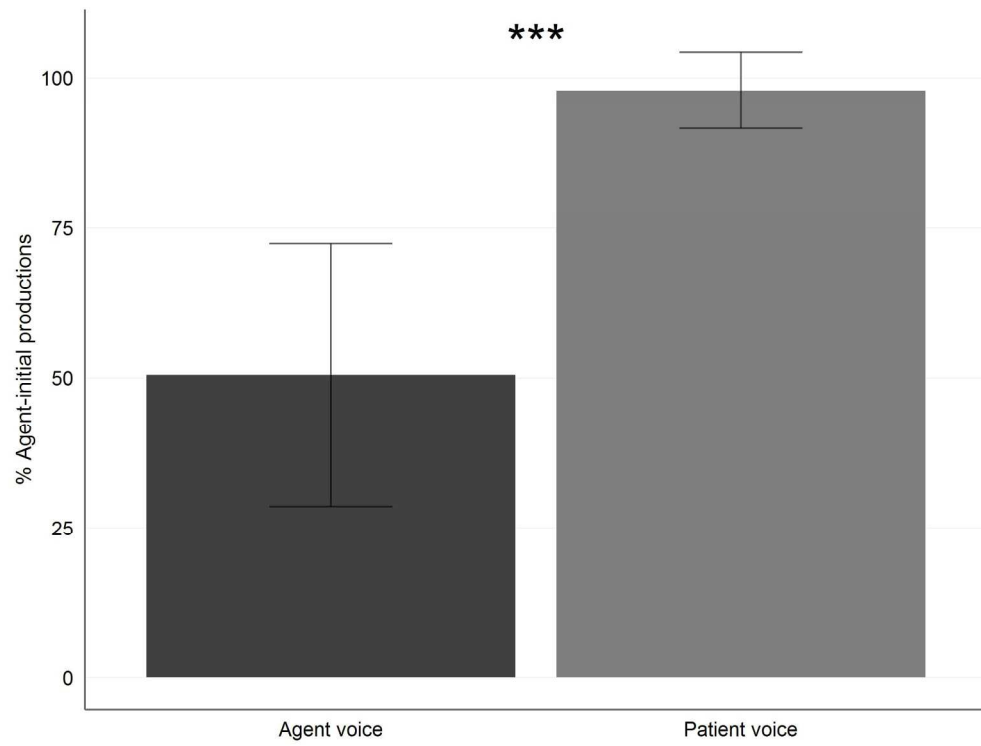


Figure 2. Mean percentage of agent-initial productions in Experiment 1 with between participant 95% confidence intervals per voice condition.

*** $p < .001$.

833x624mm (72 x 72 DPI)

view

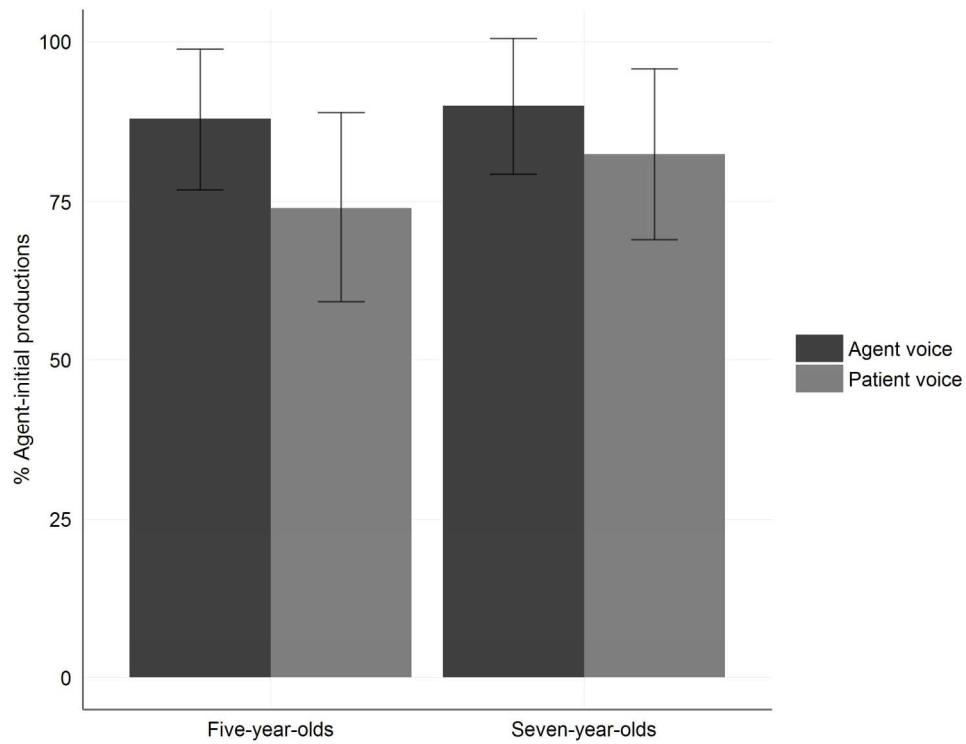


Figure 3. Mean percentage of agent-initial productions in Experiment 2 with between participant 95% confidence intervals per voice condition within each age group.

833x624mm (72 x 72 DPI)

Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

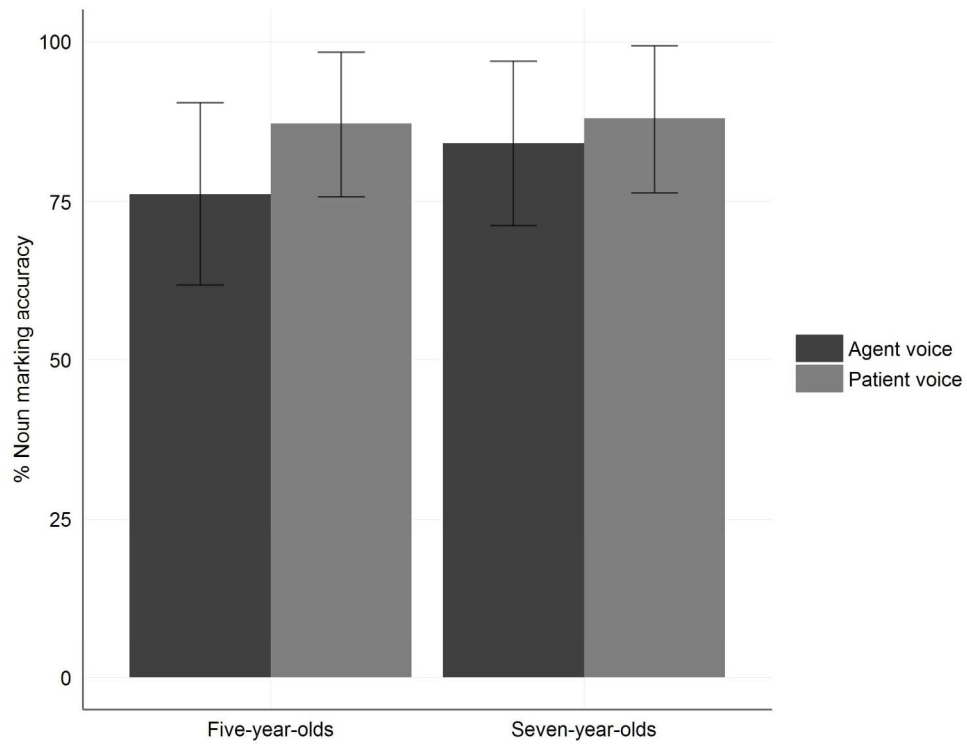


Figure 4. Mean percentage noun marking accuracy in Experiment 2 with between participant 95% confidence intervals per voice condition within each age group.

833x624mm (72 x 72 DPI)

Review

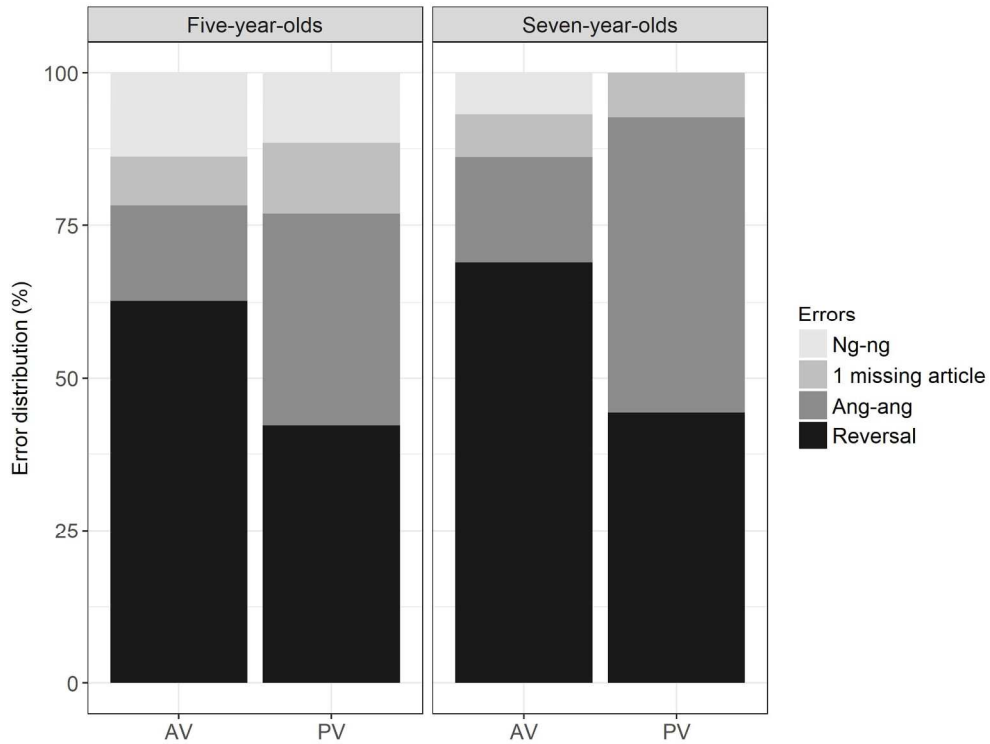


Figure 5. Percentage distribution of noun-marking error within each voice condition (AV: agent voice, PV: patient voice) within each age group in Experiment 2.

833x624mm (72 x 72 DPI)

Review