Corresponding Author:
Nozomi Tanaka, Department of East Asian Languages and Cultures, Indiana University, Global and International Studies Building 2050A, 355 N. Jordan Avenue, Bloomington, IN, 47401, USA
Email: tanakan@indiana.edu

An asymmetry in the acquisition of relative clauses

Evidence from Tagalog

Nozomi Tanaka

Indiana University

William O'Grady

University of Hawai'i at Mānoa

Kamil Deen

University of Hawai'i at Mānoa

Ivan Paul Bondoc

University of Hawai'i at Mānoa

Abstract

This paper reports on the acquisition of relative clauses in Tagalog, the most widely spoken language in the Philippines. A distinctive feature of Tagalog is a unique system of voice that creates competing patterns, each with different possibilities for relativization. Our study of children's performance on agent and patient relative clauses in a comprehension task revealed an agent relative clause advantage. These findings cannot be explained by the voice preference in declarative clauses, but are compatible with an explanation based upon input frequency factors.

Keywords

Tagalog, symmetrical voice, relative clauses, Philippine-type voice, comprehension

Introduction

A large body of research has demonstrated that not all relative clauses (RCs) are acquired and processed with equal ease (for reviews, see Diessel & Tomasello, 2005; Friedman, Belleti, & Rizzi, 2009; Kidd, Brandt, Lieven, & Tomasello, 2007; Kim & O'Grady, 2016). Typically, the asymmetry, which takes the form of a subject advantage, is observed in the comparison of subject RCs and object RCs, such as those exemplified below in English.

(1) a. Subject RC: the boy [that _ is chasing the girl]
    b. Object RC: the girl [that the boy is chasing _]

This asymmetry has been found in a number of languages, for a variety of learner populations using a variety of methodologies and has received several potential explanations. Here we briefly review two prominent explanations for this asymmetry: The Main Clause Hypothesis, and the Frequency Hypothesis.

Proponents of the Main Clause Hypothesis (Diessel & Tomasello, 2005; Diessel, 2007) argue that subject RCs in English are acquired earlier than direct object RCs because they are more similar to simple nonembedded sentences, in that the relevant nouns and verbs occur in the same order as in main clauses. They show that in an elicited imitation task, both children acquiring English and children acquiring German exhibit the same pattern of results (i.e., a subject RC preference) despite significant structural differences in the RCs of the two languages. Diessel & Tomasello attribute the similarity in the two populations to the fact that the word order in finite main clauses is the same in each language, and that this word order matches that of subject RCs. In Japanese, on the other hand, due to its SOV word order and head-final RCs, neither subject RCs ([$_{RC}$ OV] S) nor object RCs ([$_{RC}$ SV] O) resemble the basic word order (SOV). Diessel (2007) uses this observation to explain the lack of clear subject RC advantage in Japanese (Ozeki & Shirai, 2007). We generalize their insight somewhat, interpreting it to predict an advantage for RCs that most resemble the language's canonical main clauses, whose properties may involve particular inflectional patterns rather than, or in addition to, word order.

The Frequency Hypothesis, proposed by Kidd et al. (2007) and Brandt, Kidd, Lieven, and Tomasello (2009), holds that the subject RC advantage reported in many experiments is an artifact, reflecting the practice of testing children on RCs in which the direct object is animate (e.g., *the girl [that the boy pushed]*) when, in fact, there is a universal preference for inanimate direct objects (Aissen, 1999). When children are tested on the types of object RCs they hear most frequently, performance improves and the asymmetry between subject RCs and object RCs disappears.

In this paper, we present data from Tagalog, a language whose unusual system of voice and relativization distinguishes it from the languages that have been studied to date. We begin by outlining the basic morphosyntax of Tagalog. We then present a comprehension experiment that reveals an intriguing asymmetry in Tagalog relative clauses, which we associate with an interaction of factors present in the input to children.

Background on Tagalog

Tagalog manifests a system of voice that is characterized by the existence of two competing patterns for the expression of 'transitive' actions involving two participants (e.g., Foley, 1998). Each voice highlights a different argument by marking it with the 'focus marker' *ang*. As illustrated in (2), the agent voice, indicated by the infix *-um-*, highlights the verb's agent argument, while the patient voice, signaled with the infix *-in-*, has a similar effect on the patient argument.[1] The literature on Tagalog variously uses the terms 'focus,' 'subject,' 'topic,' 'pivot,' and 'privileged syntactic argument' to pick out the highlighted argument; we will adopt the term 'pivot' here for the purposes of descriptive convenience.[2]

(2) a. Agent Voice (pivot = the agent *lalake*, marked with *ang*)

H<**um**>a~habol	ang	lalake	ng	babae.
<AV>IPFV~chase	PVT	man	NPVT	woman

'The man is chasing a/the woman. '

b. Patient Voice (the pivot = the patient *babae*, marked with *ang*)

H<**in**>a~habol	ng	lalake	ang	babae.
<PV>IPFV~chase	NPVT	man	PVT	woman

'A/The man is chasing the woman. '

As illustrated in these examples, each voice is encoded by its own special verbal affix (*-um-* for agent voice and *-in-* for patient voice), neither of which is morphologically derived from the other. Moreover, neither argument in either pattern is marked as an oblique (signaled by markers such as *sa*).[3] Tagalog is therefore fundamentally unlike languages that have a canonical active transitive pattern and an alternative detransitivized construction (e.g., a passive, in which there is no direct object and the agent is encoded as an oblique argument: *The book was read by Sam*).

Another noteworthy trait of Tagalog is its verb-initial word order. The ordering of post-verbal elements is flexible, but for ease of exposition, we present our examples in verb-agent-patient word order, following the accepted consensus that this is the preferred order for children (Garcia, Dery, Roeser, Höhle, 2018; Garcia, Roeser, Höhle, 2018; Tanaka, 2016). For adults, however, there is no consensus on the canonical word order of Tagalog (see Garcia, Dery, et al., 2018 for a review). Billings (2005) argues that the canonical word order is pivot-final, while others (e.g., Aldridge, 2004) argue that the canonical order is agent-initial. In fact, it is also possible that agent voice has two canonical orders: verb-agent-patient and verb-patient-agent, as suggested by Guilfoyle et al. (1992) and Kroeger (1993). The split word order preference for agent voice is reported also in experimental studies by Garcia, Dery, et al. (2018) and Tanaka (2016). In addition to verb-initial word order, Tagalog also allows fronting of a nominal (Schachter and Otanes, 1972).

Voice plays a major role in the grammar of Tagalog by making one argument more accessible than others to various syntactic operations. Relativization is a case in point. In agent-voice patterns, relativization of the agent (3a) is fully natural, whereas relativization of the patient (3b) is unacceptable.

(3) a. Agent RC with agent voice
lalake=ng   [h<**um**>a~habol   _   ng   babae]
man=L   <AV>IPFV~chase   NPVT   woman
'(the) man [that _ is chasing a/the woman]'

   b. Patient RC with agent voice
 * babae=ng   [h<**um**>a~habol   ang   lalake _]
   woman=L   <AV>IPFV~chase   PVT   man
   '(the) woman [that the man is chasing _]'

Similarly, it is natural to relativize the patient in a patient voice clause (4a), but relativization of the patient in an agent voice clause (4b) is less acceptable (Pizarro-Guevara & Wagers, 2018) and has even been reported to be ungrammatical (Aldridge, 2004; Rackowski & Richards, 2005).

(4) a. Patient RC with patient voice
lalake=ng   [h<**in**>a~habol   ng   lalake _]
woman=L   <PV>IPFV~chase   NPVT   man
'(the) woman [that a/the man is chasing _]'

   b. Agent RC with patient voice
   ? lalake=ng   [h<**in**>a~habol   _   ang   babae]
   man=L   <PV>IPFV~chase   PVT   woman
   '(the) man [that _ is chasing the woman]'

Previous work on Tagalog has established a general preference for patient voice in basic declarative sentences. Cooreman et al. (1984) report that patient-voice patterns are used more frequently than their agent-voice counterparts in text. It is also generally reported that patient-voice patterns are mastered earlier by children than agent-voice patterns. Segalowitz and Galang (1978) conducted a picture-selection comprehension task with 30 children (10 three-year-olds, 10 five-year-olds, and 10 seven-year-olds), finding statistically significantly better performance on patient voice patterns. Along similar lines, the children in Garcia, Dery, et al.'s (2018) sentence completion task (aged 5;01–7;11) were less likely to make case marking errors (including reversal errors) when they were given a verb with a patient-voice affix. Garcia, Roeser, et al. (2018) also found in a self-paced listening task and a picture verification task that children aged 5;4–7;11 were better at using morphosyntax information in patient voice than in agent voice in terms of listening times and accuracy. Bautista (1983) reported a parallel preference in an elicited production task involving 107 Tagalog-dominant children aged 2;2 to 4;6. Tanaka (2016) also found a patient voice preference in her elicited production task, which was designed to elicit simple declarative sentences from 19 children aged 4;11 to 5;9 (mean age 5;5) as well as 9 adult controls. In situations where both arguments are animate and definite, the patient voice was used in 74.39% of the children's responses and the agent voice in just 19.51%. Moreover, even in situations that normally call for the agent voice (e.g., patterns with a definite animate agent and a non-definite inanimate patient), children used the *patient* voice in 52.87% of their responses and the agent voice in just 44.83%.

There are also indications of a patient-voice preference in the use of relative clauses. We examined the overall frequency of agent RCs and patient RCs in Marzan's (2009) Filipino Early

Language and Child Development Database (FELCDD), a corpus of child-caretaker speech to 6 children (5 female; 1 male) aged 1;2–5;0 that contained 39,120 utterances of child-directed speech. As the data summarized in Table 1 shows,[4] patient RCs are substantially more frequent than Agent RCs in speech to children.

Table 1. Frequency data on Agent RCs and Patient RCs in child-directed speech in Marzan (2009).

	Head animacy	Agent RCs	Patient RCs
Reversible	Animate	3	3
Non-reversible	Animate	20	0
	Inanimate	1	55
Total		24	58

These facts notwithstanding, some previous experimental work has shown an agent RC preference. Of particular interest is a study by Bondoc, O'Grady, Deen & Tanaka (2018) on language loss in Tagalog. They report an agent RC advantage in both comprehension and production in patients with agrammatic aphasia. We explore the possibility of a similar asymmetry in the case of children learning Tagalog, consistent with previously observed parallels between acquisition and attrition noted for other languages (e.g., Cho-Reyes & Thompson 2012 for English, and Friedmann, Reznick, Dolinsky-Nuger & Soboleva 2010 for Russian).

Method

Participants

Thirty-one Tagalog-speaking children (19 female, 12 male) aged 4;8−5;9 (mean 5;5) and 28 adult controls participated in our comprehension study. Participants were recruited and tested in Quezon City (at their school) and in Bacoor City (at a church). The child participants were all in preschool. Both child and adult participants showed varying degrees of multilingualism, as is typical of the majority of the Filipino population. All were proficient in Tagalog and had exposure to English (a second official language of the Philippines), although their English proficiency varied. In some cases, they had also been exposed to another Philippine language. However, our participant recruitment targeted Tagalog speakers (as reported by teachers and parents), and a language background survey was conducted to ensure that Tagalog is an every-day language of all the participants.

Materials

Following the methodology used in previous studies (e.g., Lau, 2016), each test item consisted of a panel of two pictures depicting reversible events (Tanaka et al., 2016). A sample test item is shown in Figure 1.
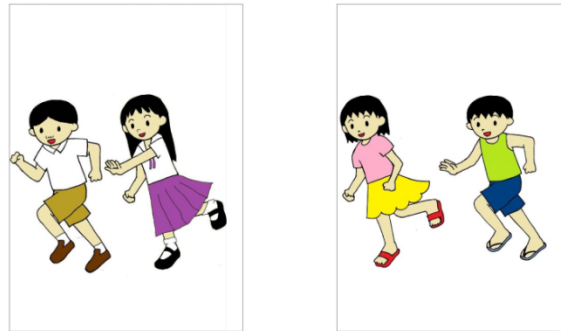

Figure 1. Sample test item with pictures of a girl chasing a boy and a boy chasing a girl.

Each of these panels contained two referents, the agent and the patient, yielding a total of four possible referents in each item. The test items depicted five events with an animate agent and an animate patient (*basa* 'splash/wet,' *buhat* 'carry,' *habol* 'chase,' *tulak* 'push,' and *yakap* 'hug').[5] These verbs appeared in two conditions: an agent RC condition and a patient RC condition.[6] In total, there were ten test items, five agent RCs and five patient RCs, which were presented in a randomized order.

Procedure

In each test trial, participants were asked to select the referent that matched the auditory description, which was given in the form of either an agent RC or a patient RC. The example in (5) illustrates an agent RC prompt for the verb *habol* 'chase,' while (6) exemplifies a patient RC prompt for the same verb.

> (5) Agent RC
> babae=ng [h<um>a~habol   ng   lalake]
> girl=L   <AV>IPFV~chase   NPVT   boy
> '(the) girl [that _ is chasing a/the boy]'

> (6) Patient RC
> babae=ng [h<in>a~habol   ng   lalake]
> girl=L   <PV>IPFV~chase   NPVT   boy
> '(the) girl [that a/the boy is chasing _]'

Adults made their choice with the help of a MouseTracker (Freeman & Ambady, 2010). Children were instructed to point to the character of their choice as a research assistant recorded the location of their finger. The test session began with a brief practice session, during which participants were trained to point to the character of their choice rather than the whole picture in which it occurred. The practice session involved four items with non-reversible relative clauses (e.g., 'the girl that is picking the flower' vs. 'the girl that is picking the tomato'; 'the flower that

6

the boy is picking' vs. 'the flower that the girl is picking'). Figure 2 presents a sample practice item.
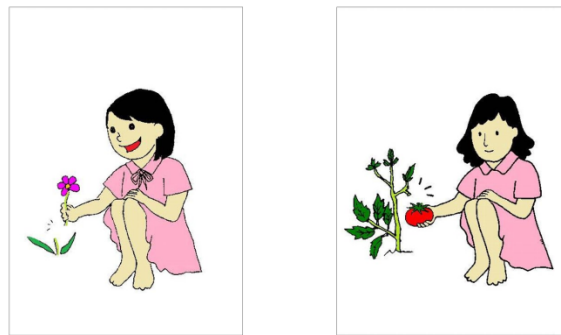

Figure 2. Sample practice item for an agent RC 'a girl that is picking the tomato.'

Although the practice items were themselves RCs, they crucially involved non-reversible events, and so were significantly easier to interpret than the actual test items.

Analysis

Participants' responses were classified with the help of the answer types illustrated in Figure 3.
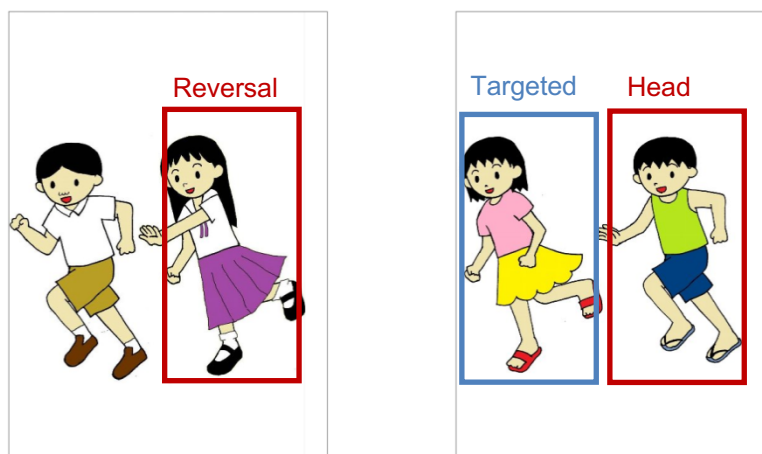

Figure 3. Response types for a patient RC 'the girl that the boy is chasing.'

Given the patient RC 'the girl that the boy is chasing,' the correct answer is to pick the girl in the right-hand picture—the 'targeted response.' We treated selection of the girl in the picture on the left as a reversal error since it depicts the girl that is chasing the boy rather than vice versa. The selection of the boy in the correct picture was coded as a head error since it depicts the wrong type of person (a boy rather than a girl). Other erroneous responses include choosing the boy in the wrong picture and clicking on (or pointing at) the white space in the picture.

Results

Two adult participants had 0% accuracy in the patient RC condition, and thus were removed from the analysis.[7] Table 2 summarizes the accuracy rates for adults and children. Adults performed at ceiling for both types of RCs (96.92% for agent RC and 95.38% for patient RCs). Children's mean accuracy rate for agent RCs was 49.03%, compared to 14.84% for patient RCs.[8]

Table 2. Adults' and children's results in the comprehension task. (Means are calculated by participants.)

	Adults				Children			
	Agent RC		Patient RC		Agent RC		Patient RC	
	Token	%	Token	%	Token	%	Token	%
Targeted	126	96.92	124	95.38	76	49.03	23	14.84
Reversal error	3	2.31	3	2.31	18	11.61	56	36.13
Head errors	1	0.77	1	0.77	41	26.45	40	25.81
Other errors[9]	0	0.00	2	1.54	20	12.90	36	23.23
Total	130	100	130	100	155	100	155	100

The data were fitted into a binomial mixed-effects model using R (R Core Team, 2017) and lme4 package (Bates, Maechler, Bolker, & Walker, 2015). The model included participant group (adults vs. children) and RC type (Agent RC vs. Patient RC) as fixed effects, as well as the interaction between the two. As random effects, the model included participants, items, and trial order as random intercepts, as well as by-participant effects and by-item effects of RC type as a random slope. We used type II Wald χ^2 tests to test the significance of the fixed effects and report the results in Table 3.

Table 3. Results of the binomial mixed effects model for the comprehension task.

	Estimate (β)	Standard Error (SE)	χ2	p
Intercept	4.12	0.69		
Group (children)	−4.18	0.74	76.27	< .001 ***
RC type (Patient RC)	−0.54	0.98	10.60	.001 ***
Group × RC type	−1.62	1.16	1.93	.164

Note: Model: Accuracy ~ Group * RC type + (1 + RC type | Participant) + (1 + RC type | Item) + (1 | Trial order)
p-values based on type II Wald χ^2 tests. Significance levels: * *p* < .05, ** *p* < .01, *** *p* < .001

There was a significant main effect of group ($p < .001$). A pairwise comparison using estimated marginal means showed no significant difference between RC types for adults ($p = .583$) but a significant difference between RC types for children ($p = .001$). In other words, whereas adults showed no preference for agent RCs or patient RCs, children were significantly more accurate in their comprehension of agent RCs.

The most commonly attested errors involved reversals. Consistent with the existence of an agent RC preference, children interpreted patient RCs as agent RCs (a reversal error) about three times more often than vice versa (children: agent RC 11.61%, patient RC 36.13%). This difference was significant ($p = .002$). The asymmetrical nature of these errors (patient RCs interpreted as agent RCs, but not agent RCs interpreted as patient RCs) is a further indication of an agent RC advantage.

Discussion

While adult performance did not significantly differ between the two types of RCs, children showed better accuracy in comprehending agent RCs than patient RCs. Moreover, the distribution of reversal errors indicates that they were significantly more likely to interpret patient RCs as agent RCs than the other way around. These findings point toward an agent RC advantage in the comprehension of RCs. Why should this be the case?

It is clear that the agent RC preference cannot be attributed to the relative frequency of agent voice patterns in the input, as predicted by the Main Clause Hypothesis, even if we generalize this to mean that preferred main clauses should be identified based on voice marking and case rather than word order.

We examined the frequency of agent and patient voice patterns for each of the five verbs used in our experiment, basing our calculations on McFarland's (1989) lexical frequency data (from written texts) and on the child-directed data from Marzan (2009). As illustrated in Table 4, there is no significant asymmetry for two of the verbs that we used (*basa* 'wet' and *buhat* 'carry'), and the patient voice form is dominant for the remaining three verbs.

Table 4. Frequency data on the 10 verbs used in the comprehension experiment.

	McFarland (1989)		Marzan (2009)	
	agent voice	patient voice	agent voice	patient voice
basa 'wet'	-	-	-	-
buhat 'carry'	-	-	1	1
habol 'chase'	27.0	54.2	3	14
tulak 'push'	13.4	19.1	-	3
yakap 'hug'	11.4	36.6	-	-

Note: The numbers from McFarland's study reflect normalized values, whereas those from Marzan's studies represent a raw token count for clauses with two overt arguments.

Therefore, if we take Diessel and Tomasello's Main Clause Hypothesis to mean that an advantage for RCs stems from the particular inflectional patterns involved in a language's canonical main clauses, it would predict a patient RC advantage based on the predominance of PV. However, this was not the case.

If we take both word order and voice into account, the Main Clause Hypothesis would predict no difference between agent and patient RCs because agent- and patient-initial sentences are both non-canonical even though they are possible and frequent word order patterns (Table 5; Cooreman et al., 1984).

Table 5. Voice and word order frequency in Cooreman et al. (1984, p. 17)

	Frequency	Percentage
AV, verb-initial	37	13%
AV, agent-initial	31	11%
PV, verb-initial	166	59%
PV, patient-initial	47	17%
Total	281	100%

However, there is a somewhat different account of how an agent RC advantage might emerge as the result of the interaction between RC type and animacy of the patient argument, consistent with the Frequency Hypothesis. The key observation is that the animacy of the head noun appears to correlate strongly with the use of agent voice in relative clauses.[11] As shown in

Table 1, of the 26 RC patterns with animate heads in maternal speech (the first two rows of the table), 23 (88.46%) are agent RCs; just 3 (11.53%) are patient RCs. Moreover, of the 56 RCs with inanimate heads, just 1 (1.81%) involves an agent RC. Thus, exposure to a relative clause pattern that begins with an animate head noun should trigger in the child a very strong expectation that the RC will be of the agent type—an expectation that fits well the high success rate on agent RCs compared to their patient RC counterparts.

While the data we presented are of the most relevance to the Main Clause Hypothesis and the Frequency Hypothesis, other explanations for the asymmetry also warrant consideration.

There have been analyses that focus on structural features of the favored pattern. In the distance-based theories (e.g., Wanner and Maratsos, 1978; Gibson, 1998, 2000), a subject relative clause has smaller working memory cost than an object relative clause, as the gap in a subject relative clause is closer to the relativized element than the gap in an object relative clause. However, it is not clear that Tagalog offers direct insights into these theories. For example, as we have previously explained, there is no consensus on whether the 'basic' order within relative clause is verb-agent-patient or verb-patient-agent, since both are allowed. We therefore cannot say with certainty whether the gap in an agent RC is closer to the relativized element than the gap in a patient RC. This precludes the possibility of a clear-cut prediction by distance-based theories of RC difficulty (e.g., Wanner and Maratsos, 1978; Gibson 1998, 2000).

(7)   Possible positions of the gap in an agent relative clause.
                              here   *or here*
                                 ↓   ↓
        babae=ng   [h<um>a~habol   ng   lalake   ]
        girl=L   <AV>IPFV~chase   NPVT   boy
         '(the) girl [that _ is chasing a/the boy]'

Other explanations make reference to the position of the relativized element in syntactic structure (e.g., Collins, 1994; Friedman et al., 2009). In such explanations, the subject relative clause advantage manifests because the gap in a subject relative clause is less deeply embedded in syntactic structure than the gap in an object relative clause (e.g., Collins, 1994), or because there is no intervening element between the head and the gap in a subject relative clause unlike in an object relative clause (Friedmann et al., 2009). While the syntactic structure of Tagalog remains in dispute, in some of the well-known theories (Aldridge, 2012; Rackowski & Richards, 2005), an argument that is to be relativized moves to the 'subject' position before relativization occurs—which would mean that both the agent and the patient would originate in the same structural position in their respective RCs. If correct, this would rule out various explanations for RC difficulty that make reference to the position of the relativized element in syntactic structure (e.g., Collins, 1994; Friedman et al., 2009).

Lastly, we cannot rule out possible semantic and pragmatic explanations. For example, it has been suggested that, all other things being equal, agents are more salient and more topical than other arguments, by virtue of their role as instigators of the event (e.g., Cohen & Oakes, 1993; Kim & O'Grady, 2016).[10] This opens another possible line of explanation for the agent RC advantage which might well be worth exploring in future research.

Conclusion

This study investigated the acquisition of relative clauses in child Tagalog. We found that children are significantly more accurate at comprehending agent RCs than patient RCs and are more likely to convert patient RCs into agent RCs than the other way around. This agent RC advantage exists despite the fact that the preferred verbal inflection in the language in general involves the patient voice, and despite the fact that patient RCs are overall more common than agent RCs in maternal speech. The key factor underlying children's interpretation of RCs, we suggest, lies in the interaction between the animacy of the head noun and its strong correlation with agent RCs in the input—a correlation which in turn can be traced to the strong correlation between animacy and agency (Dowty, 1991).

      Our study is one of the first to offer data showing an asymmetry in the acquisition of RCs in a language with the type of voice-based syntax that characterizes Tagalog. Our findings highlight the importance of work on the acquisition of relative clauses to a broader understanding of development, making their further study all the more intriguing and urgent. Hopefully, additional study of languages with less common systems of relativization will shed further light on this matter.

Acknowledgements

Funding

Notes

1. Following Leipzig Glossing Rules, angle brackets (<>) enclose infixes, and a tilde (~) indicates reduplication. List of abbreviations: AV = agent voice, IPFV = imperfective, PVT = pivot, L =linker, NPVT = non-pivot, PV = patient voice, RC = relative clause.
2. Various analyses have been proposed regarding the nature of the pivot. Some researchers have argued that it is a subject (Guilfoyle, Hung, & Travis, 1992), while others have proposed that it signals a sentential topic (Schachter & Otanes, 1972), a discourse topic (Cooreman, Fox, & Givón, 1984; Katagiri, 2006), or an absolutive (Aldridge, 2004). Moreover, the selection of pivot is known to be influenced by different factors such as definiteness (Foley & Van Valin, 1984; Reid & Liao, 2004), specificity (Maclachlan &

Nakamura, 1997; Rackowski & Richards, 2005), genericity, referentiality, topicality (Carrier-Duncan, 1985), and agentivity (Saclot, 2006).

3. While *sa* generally markes locatives or datives, the non-pivot agent in patient voice can be marked as *sa* if it is human and specific (Himmelman, 2005).

4. The data include all possible RC patterns in Tagalog (head-initial, head-final, head-internal, and headless RCs; see Aldridge, 2004 and Law, 2015) as well as data from *wh*-questions and pseudo-cleft constructions, as they contain free relatives. There were 1 agent RC and 9 patient RCs that contained the head.

5. Following the usual practice, we did not attempt to test children's comprehension of non-reversible RCs (e.g., 'the book that the boy read'), whose correct interpretation can be inferred without attention to morpho-syntactic cues.

6. This means that each participant encountered each verb two times across two conditions. There was no obvious sign of learning effects, in which the participants were more successful with the second appearance of the verb (35.33%) than the first appearance of the verb (30.67%). An anonymous reviewer suggested a possibility that a participant gave the same answer for the two instances of one verb. When we focus on the verb forms and disregard response types, children gave the same interpretation to both occurrences of the verb 76.12% of the time. However, 51.61% of them were agent relative clause interpretation, therefore this is likely the byproduct of the agent relative clause advantage.

7. Inclusion of these two adult participants in the analysis did not change the overall results.

8. Accuracy rate for each verb is as follows: *yakap* 'hug' 35.48%, *basa* 'wet' 33.87%, *buhat* 'carry' 32.36%, *tulak* 'push' 30.65%, *habol* 'chase' 27.42%. There was no obvious correlation between the accuracy rate and verb frequency. There was no clear trend in which each verb is strongly associated with one voice or the other.

9. The "other" responses from adults in the patient RC condition were uninterpretable based on the location of the clicking. All of the "other" responses from children involved the selection of the wrong head in the wrong picture (e.g., 'the boy that the girl is chasing' instead of 'the girl that the boy is chasing'.)

10. Indeed, animacy strongly correlates with agent voice in all transitive patterns in the language, including basic declarative clauses. Tanaka (2016) reports that adults produced agent voice 70% of the time when the event involves an animate agent and an inanimate patient, but when the agent and the patient are both animate, adults produced patient voice without an exception.

11. There are also possible language-specific effects at play. Sauppe et al. (2017) measured the pupil size of adult Tagalog speakers during a picture description task and reported that the production of PV is more cognitively taxing for speakers, even though they produced more instances of PV than AV. We thank an anonymous reviewer for suggesting this possibility.

References

Aissen, J. (1999). Markedness and subject choice in Optimality Theory. *Natural Language & Linguistic Theory*, *17*, 673–711.

Aldridge, E. (2004). Internally headed relative clauses in Austronesian languages. *Language and Linguistics*, *5*(1), 99–129.

Aldridge, E. (2012). Antipassive and ergativity in Tagalog. *Lingua*, *122*(3), 192–203.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software 67*, 1–48.

Bautista, M. L. S. (1983). Word order in Tagalog child language. *Philippine Journal of Linguistics 14*, 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* (pp. 303–340). Amsterdam/Philadelphia: John Benjamins.

Bondoc, I. P., O'Grady, W., Deen, K., & Tanaka, N. (2018). Agrammatism in Tagalog: voice and relativisation. *Aphasiology*, *32*(5), 598–617.

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: an investigation of young German and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics*, *20*, 539–70.

Carrier-Duncan, J. (1985). Linking of thematic roles in derivational word formation. *Linguistic Inquiry*, *16*(1), 1–34.

Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production and comprehension in aphasia: Northwestern assessment of verbs and sentences (NAVS). *Aphasiology*, *26*, 1250–1277.

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, *29*(3), 421–433.

Collins, C. (1994). Economy of derivation and the generalized proper binding condition. *Linguistic Inquiry*, *25*(1), 45–61.

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

Diessel, H. (2007). A construction-based analysis of the acquisition of East Asian relative clauses. *Studies in Second Language Acquisition*, *29*, 311–320.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, *81*, 882–906.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619.

Foley, W. A. (1998, July 1). Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the Workshop on Voice and Grammatical Functions in Austronesian Languages, Lexical Functional Grammar 98 Conference, University of Queensland in Brisbane, Australia.

Foley, W. A., & Van Valin, R. D. (1984). *Functional syntax and universal grammar.* Cambridge, UK: Cambridge University Press.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*, 67–88.

Friedmann, N., Reznick, J., Dolinski-Nuger, D., Soboleva, K. (2010). Comprehension and production of movement-derived sentences in Russian speakers with agrammatic aphasia. *Journal of Neurolinguistics*, *23*, 44–65.

Garcia, R., Dery, J. E., Roeser, J., & Höhle, B. (2018). Word order preferences of Tagalog-speaking adults and children. *First Language*, *38*(6), 617–640.

Garcia, R., Roeser, J., & Höhle, B. (2018). Thematic role assignment in the L1 acquisition of Tagalog: Use of word order and morphosyntactic markers. *Language Acquisition*.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Guilfoyle, E., Hung, H., & Travis, L. (1992). SPEC of IP and SPEC of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory*, *10*(3), 375–414.

Himmelmann, N. P. (2005). Tagalog. In K. A. Adelaar & N. P. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 350–376). London: Routledge.

Katagiri, M. (2006, January). *Topichood of the Philippine topic: Revisited from a cross-linguistic perspective*. Paper presented at the 10th International Conference on Austronesian Linguistics, Palawan, Philippines. Retrieved from https://www.sil.org/resources/archives/25584

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: a cross-linguistic comparison of the contraints influencing young children's processing of relative clauses. *Language and Cognitive Processes 22*, 860–97.

Kim, C.-E., & O'Grady, W. (2016). Asymmetries in children's production of relative clauses: Data from English and Korean. *Journal of Child Language*, *43*, 1038–71.

Kroeger, P. R. (1993). Another look at subjecthood in Tagalog. *Philippine Journal of Linguistics*, *24*(2), 1–16.

Lau, E. (2016). The role of resumptive pronouns in Cantonese relative clause acquisition. *First Language*, *36*(4), 355–382.

Law, P. (2015). The syntax of Tagalog relative clauses. In A. Camp, Y. Otsuka, C. Stabile, & N. Tanaka (Eds), *AFLA 21: The Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association* (pp. 167–184). Canberra: Asia-Pacific Linguistics. Retrieved from http://hdl.handle.net/1885/95329

Maclachlan, A., & Nakamura, M. (1997). Case-checking and specificity in Tagalog. *The Linguistics Review*, *14*, 307–333.

Marzan, J. C. B. (2009). Filipino Early Language and Child Development Database (FELCDD): report. Unpublished manuscript, University of the Philippines Diliman.

McFarland, C. D. (1989). *A frequency count of Pilipino*. Manila, Philippines: Linguistic Society of the Philippines.

Ozeki, H., & Shirai, Y. (2007). Does the Noun Phrase Accessibility Hierarchy predict the difficulty order in the acquisition of Japanese relative clauses? *Studies in Second Language Acquisition*, *29*(02), 169–196.

Pizarro-Guevara, J., & Wagers, M. (2018, May). *Agent extraction under patient voice in Tagalog is acceptable: Evidence from acceptability ratings in various A'-dependencies*. Presented at the 25th Meeting of the Austronesian Formal Linguistics Association, Taipei, Taiwan.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Rackowski, A., & Richards, N. (2005). Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry*, *36*(4), 565–599.

Reid, L. A., & Liao, H. (2004). A brief syntactic typology of Philippine languages. *Language and Linguistics*, *5*(2), 433–490.

Saclot, M. J. D. (2006). *On the Transitivity of the Actor Focus and Patient Focus Constructions in Tagalog*. Paper presented at the 10th International Conference on Austronesian

Linguistics, Palawan, Philippines. Retrieved from
https://www.sil.org/resources/publications/entry/25914

Sauppe, S. (2017). Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language*, *93*(2), 288–313.

Schachter, P., & Otanes, F. T. (1972). *Tagalog reference grammar*. Berkeley: University of California Press.

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

Tanaka, N. (2016). *An asymmetry in the acquisition of Tagalog relative clauses*. Ph.D. dissertation, University of Hawai'i at Manoa.

Tanaka, N., O'Grady, W., Deen, K., Kim, C.-E., Hattori, R., Bondoc, I. P. M., & Soriano, J. U. (2016). "Relative clause comprehension task." Nozomi Tanaka Collection. Kaipuleohone: http://hdl.handle.net/10125/42194. Type: language description. Media: image. Access: public. Resource ID: NT1-003.

Wanner, E., & Marastos, M. (1978). An ATN approach to comprehension. In J. Bresnan, G. A. Miller, & M. Halle (Eds.), *Linguistic Theory and Psychological Reality* (pp. 119–161).