More Relativization Asymmetries:
Children Find Locative and Benefactive Clauses Difficult

Ivan Paul Bondoc, William O'Grady, Kamil Ud Deen, Nozomi Tanaka, Ellyn Cassey Chua, Angela Claudine De Leon, and Joshua Arvin Siscar

1. Background

The syntax of Tagalog is characterized by a system of voice alternations that highlights one of the verb's arguments (e.g., an agent, theme, benefactive, or locative) with the help of verbal affixation and the use of the special case marker *aŋ*. The examples in (1) illustrate agent-voice and theme-voice patterns, both of which are frequently attested in Tagalog speech.

(1) a. Agent Voice (AV) Declarative
         Nag-luto   aŋ   babae   para   sa   lalaki.[1]
         AV.PRF-cook   PIV   girl   for   OBL   boy
         'The girl cooked for the boy.'

     b. Theme Voice (TV) Declarative
         I-p\<in\>atong   aŋ   ruler   sa   notebook.
         TV-\<PRF\>place   PIV   ruler   OBL   notebook
         '(Someone) placed the ruler on the notebook.'

---

[1] Abbreviations: ARC = agent relative clause, AV = agent voice, BRC = benefactive relative clause, BV = benefactive voice, DEM = demonstrative, EXIST = existential, LRC = locative relative clause, LV = locative voice, NPIV = non-pivot, OBL = oblique, PIV = pivot, PRF = perfective, RC = relative clause, REL = relativizer, STATIVE = stative, TRC = theme relative clause, TV = theme voice

Less common, but still fully natural, are the two patterns illustrated in (2), which highlight the benefactive argument and the locative argument, respectively.

(2) a. Benefactive Voice (BV) Declarative
      Ip<in>ag-luto   naŋ babae aŋ   lalaki.
      <PRF>BF-cook   NPIV girl   PIV   boy
      'The girl cooked for the boy.'

    b. Locative Voice (LV) Declarative
      P<in>atung-an   naŋ ruler aŋ   notebook.
      <PRF>place-LV   NPIV ruler   PIV   notebook
      '(Someone) placed the ruler on the notebook.'

The highlighted argument, often called the 'pivot,' is exclusively eligible to undergo various syntactic operations, including relativization. As the following examples help illustrate, relativization of an agent argument calls for the agent voice, relativization of the theme argument requires the theme voice, and so on.

(3) a. Agent Relative Clause (ARC) with AV
      babae=ŋ   nag-luto   para sa   lalaki
      girl=REL   AV.PRF-cook   for OBL boy
      'the girl who cooked for the boy'

    b. Theme Relative Clause (TRC) with TV
      ruler   na   i-p<in>atong   sa   notebook
      ruler   REL TV-<PRF>place   OBL notebook
      'the ruler that (someone) placed on the notebook'

    c. Benefactive Relative Clause (BRC) with BV
      lalaki=ŋ   ip<in>ag-luto   naŋ   babae
      boy=REL   BV<PRF>-cook   NPIV girl
      'the boy who the girl cooked for'

    d. Locative Relative Clause (LRC) with LV
      notebook   na p<in>atung-an   naŋ ruler
      notebook   REL <PRF>place-LV   NPIV ruler
      'the notebook where (someone) placed the ruler on'

The typological rarity of this type of relativization as well as its morphosyntactic intricacy raise obvious questions about the manner in which it emerges in the course of linguistic development. Recent work in the acquisition of relative clauses in Tagalog has shown that children perform better on agent relative clauses compared to theme relative clauses (Tanaka, 2016), but no work has yet examined other types of relative clause. This study seeks to remedy this

situation by exploring the performance of native Tagalog children on benefactive and locative relative clauses.

2. Current Study

We conducted two experiments: an elicited production task and a character selection task designed to assess comprehension. In order to focus on the morphosyntactic contrasts among the four patterns, it was necessary to control for other factors, including the number of arguments per clause and the possibility of semantic biases.

All of our test items contained two arguments, one of which was relativized. In addition, each argument had a prototypical referent for its thematic role: agents and benefactives were always animate, and themes and locatives were always inanimate. Finally, each clause was reversible, in the sense that the two arguments could be transposed without affecting the naturalness of the sentence. This allowed us to make two comparisons based purely on morphosyntactic differences among RC types—one between agent and benefactive RCs, both of which relativize animate arguments (see (3a) and (3c) above), and one between theme and locative RCs, both of which relativize inanimate arguments ((3b) and (3d)).

2.1. Norming Study

All declarative and relative clauses used in the experiments were normed using a six-point acceptability judgment task (AJT) conducted online. Participants were asked to rate the naturalness of declarative and relative clauses as descriptions of events in corresponding pictures, with 1 being "very unnatural" and 6 being "very natural."

2.1.1. Declarative Clauses

Twenty-two native Tagalog adults (age range 20–58; mean = 28.77) participated in the norming of declarative clauses. The items included 6 AV, 6 BV, 7 TV, and 7 LV sentences that accurately described the accompanying picture, and 6 AV, 6 BV, 7 TV, and 7 LV sentences that did not. An additional 40 declarative sentences, representing a variety of different voice patterns, were included as fillers. All items in the survey were randomized. A sample of a declarative clause item normed is shown in Figure 1.

The results showed high mean, median, and mode scores for the declarative clauses that accurately described the pictures, and low mean, median, and mode scores for those that did not (Table 1).
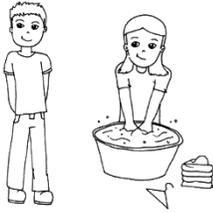
Figure 1. Sample declarative clause item (BV clause) in the norming study. Sentence: *Ipinaglaba naŋ babae aŋ lalaki.* 'The girl washed for the boy.'
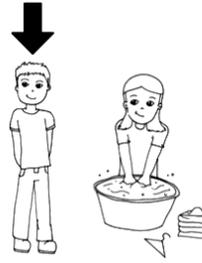
Figure 2. Sample relative clause item (BRC) in the norming study. RC: *lalakiŋ ipinaglaba naŋ babae* 'the boy for whom the girl washed'

Table 1. Results for the declarative clause items in the norming study

Sentence Type	Accuracy of description	mean	median	mode
AV	accurate AV	5.61	6	6
	inaccurate AV	1.55	1	1
BV	accurate BV	5.20	6	6
	inaccurate BV	1.76	1	1
TV	accurate TV	5.36	6	6
	inaccurate TV	1.52	1	1
LV	accurate LV	4.92	6	6
	inaccurate LV	1.61	1	1

2.1.2. Relative Clauses

Twenty adult native speakers of Tagalog (age range 23–69; mean = 29.8) participated in the norming of relative clauses. The items included 6 ARCs, 6 BRCs, 7 TRCs, and 7 LRCs that correctly described a designated character/item in the accompanying picture, and 6 ARCs, 6 BRCs, 7 TRCs, and 7 LRCs that did not. The AJT also contained 109 relative clauses of various types as fillers. All items in the survey were randomized. A sample of a relative clause item from the norming task is shown in Figure 2.

Table 2. Results for the relative clause items in the norming study

RC Type	Accuracy of description	mean	median	mode
ARC	accurate ARC	5.87	6	6
	inaccurate ARC	1.71	1	1
BRC	accurate BRC	5.10	6	6
	inaccurate BRC	1.16	1	1
TRC	accurate TRC	5.54	6	6
	inaccurate TRC	1.43	1	1
LRC	accurate LRC	5.11	6	6
	inaccurate LRC	1.43	1	1

The results showed high mean, median, and mode scores for the relative clauses that accurately described the pictures, and low mean, median, and mode scores for those that did not (Table 2). Overall, the results of the norming study confirm the appropriateness of the declarative and relative clauses to be used for the study.

2.2. Experiment 1: Relative Clause Elicited Production Task
2.2.1. Participants

Nineteen adults (age range 19–67, mean = 42.26) and eight children (age range 6;4–7;2, mean = 6;10) participated in the relative clause elicited production task.

2.2.2. Method

Following Hsu, Hermon, and Zuckowski (2009), Kim & O'Grady (2015), and Tanaka (2016), participants were shown a panel containing two pictures, as in Figures 3 and 4. An audio-recorded prompt described the event depicted in each picture. An arrow then appeared over one of the characters or items, which the participant was asked to describe with the help of a prompt such as 'Who/what has the arrow?'
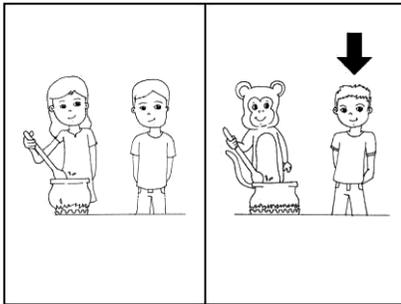

Figure 3. Sample BRC production item. Prompt (in AV): *Nagluto aŋ babae para sa lalaki. Nagluto aŋ unggoy para sa isa paŋ lalaki. Sino aŋ may arrow?* 'The girl cooked for the boy. The monkey cooked for another boy. Who has the arrow?' Target (BRC): *lalakiŋ ipinagluto naŋ unggoy* 'the boy who the monkey cooked for'
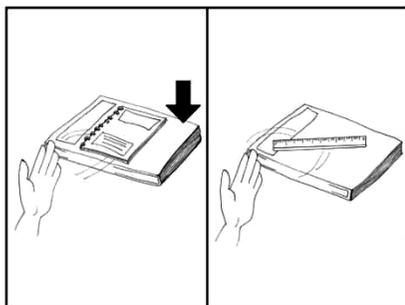
Figure 4. Sample LRC production item. Prompt (in TV): *Ipinatoŋ aŋ notebook sa libro. Ipinatoŋ aŋ ruler sa isa paŋ libro. Alin aŋ may arrow?* '(Someone) placed the notebook on the book. (Someone) placed the ruler on another book. Which one has the arrow?' Target (LRC): *libroŋ pinatungan naŋ notebook* 'the book which (someone) placed the notebook on'

The voice used in the prompt was systematically varied so as to account for possible priming effects. In one set of test items, the voice in the prompt matched the voice of the targeted pattern; in another set, it differed (e.g., agent voice for a pattern that targeted the benefactive voice and vice versa). Four lists varying in item order and voice of the prompt sentences were prepared. Each list had 3 practice items that involved intransitives, a set of 12 ARC and BRC pairs, and another set of 14 TRC and LRC pairs. Participants were pseudo-randomly assigned to one of these lists (Table 3).

Table 3. Lists used for the relative clause elicited production task.

	Voice of the Prompt Sentence	
List	for the ARC/BRC pair	for the TRC/LRC pair
1	AV	TV
2	AV	LV
3	BV	TV
4	BV	LV

2.2.3. Results

Six adult participants who did not produce relative clauses on more than half of the stimuli items were excluded, on the grounds that they had apparently not understood the protocol of the experiment. This left thirteen adult participants (22–67, mean = 44.46) for the analysis.

All participant responses were transcribed and classified into categories. All grammatically acceptable verbal RCs were coded as grammatical RCs, and were further analyzed as target responses if they included the target head noun and the target voice affixation on the verb.

(4) Sample LRC target response for Figure 4
    libro=ŋ   p<in>atung-an   naŋ   notebook
    book=REL   <PRF>place-LV   NPIV   notebook
    'the book where (someone) placed the notebook on'

Grammatical RCs that included the correct target head noun along with a stative verb were classified as non-target acceptable responses.

(5) Sample LRC acceptable response for Figure 4
    libro=ŋ   may   naka-patong   na notebook
    book=REL   EXIST   STATIVE-place   REL notebook
    'the book where there is a notebook placed on (it)'

Erroneous responses involved reversal errors, head errors, and resumptive NP errors. In a *reversal error*, the correct head is chosen, but the thematic roles are switched. A *head error* has a wrong argument as the head, while a *resumptive NP error* has the gap of the RC filled by a copy of the head noun or by a coreferential pronominal.

(6) Sample LRC reversal error
    *libro=ŋ   i-p<in>atong   naŋ   notebook
     book=REL   TV-<PRF>place-LV   NPIV   notebook
    'the book that (someone) placed on the notebook'

(7) Sample LRC head error
    notebook   na p<in>atung-an   naŋ   libro
    notebook   REL <PRF>place-LV   NPIV   book
    'the notebook that (someone) placed the book on'

(8) Sample LRC resumptive NP error
    *libro=ŋ   p<in>atung-an   ito   naŋ   notebook
     book=REL   <PRF>place-LV   DEM   NPIV   notebook
    'the book that (someone) placed the notebook on it'

Overall, the relative clause elicited production task revealed that both adults and children produced a higher number of grammatical RC responses in the ARC condition compared to the BRC condition, and in the TRC condition than in the LRC condition; see Tables 4 and 5.[2] Most errors involved reversals, i.e., the use of wrong voice affixation or case marking which led to the switching of one relative clause type to another.

---

[2] The higher number of target responses by adults on LRCs compared to TRCs may reflect the availability of more alternative strategies for describing the test picture in the latter case.

Table 4. Accuracy rates for adult participants ($n = 13$) in the relative clause elicited production task.

Response Type	*Animate Heads*		*Inanimate Heads*	
	ARC	BRC	TRC	LRC
Grammatical RCs	**90.91%**	**83.12%**	**95.56%**	**93.10%**
Target	90.91%	83.12%	51.11%	72.41%
Acceptable	0.00%	0.00%	44.44%	20.69%
Errors	9.09%	16.88%	4.44%	6.90%
Reversal	7.79%	12.99%	2.22%	4.60%
Head	1.30%	0.00%	1.11%	1.15%
Resumptive	0.00%	2.60%	1.11%	1.15%
Other[3]	0.00%	1.30%	0.00%	0.00%

Table 5. Accuracy rates for child participants ($n = 8$) in the relative clause elicited production task.

Response Type	*Animate Heads*		*Inanimate Heads*	
	ARC	BRC	TRC	LRC
Grammatical RCs	**62.22%**	**32.43%**	**64.15%**	**37.78%**
Target	62.22%	29.73%	41.51%	28.89%
Acceptable	0.00%	2.70%	22.64%	8.89%
Errors	37.78%	67.57%	35.85%	62.22%
Reversal	28.89%	45.95%	28.30%	48.89%
Head	0.00%	0.00%	0.00%	8.89%
Resumptive	6.67%	21.62%	3.77%	4.44%
Other	2.22%	0.00%	3.77%	0.00%

The elicited production data were fitted to a mixed effects logistic regression model, using the *glmer* function of the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017) in R (R Core Team, 2013). All responses were coded as either a grammatical RC or not, using binary coding. Following Barr, Levy, Scheepers, & Tily (2013), all analyses began with the full model, including RC type and prompt voice as fixed effects, and participants and verbs as random effects. The full model also included intercepts and random slopes with both RC type and prompt. All predictors were centered using contrast coding. Terms were removed one by one if the model did not converge; a likelihood ratio test was performed for model comparisons, as needed.

---

[3] The category "Other" includes wrong verb choice, presence of an additional conjunction, or a combination of the aforementioned errors.
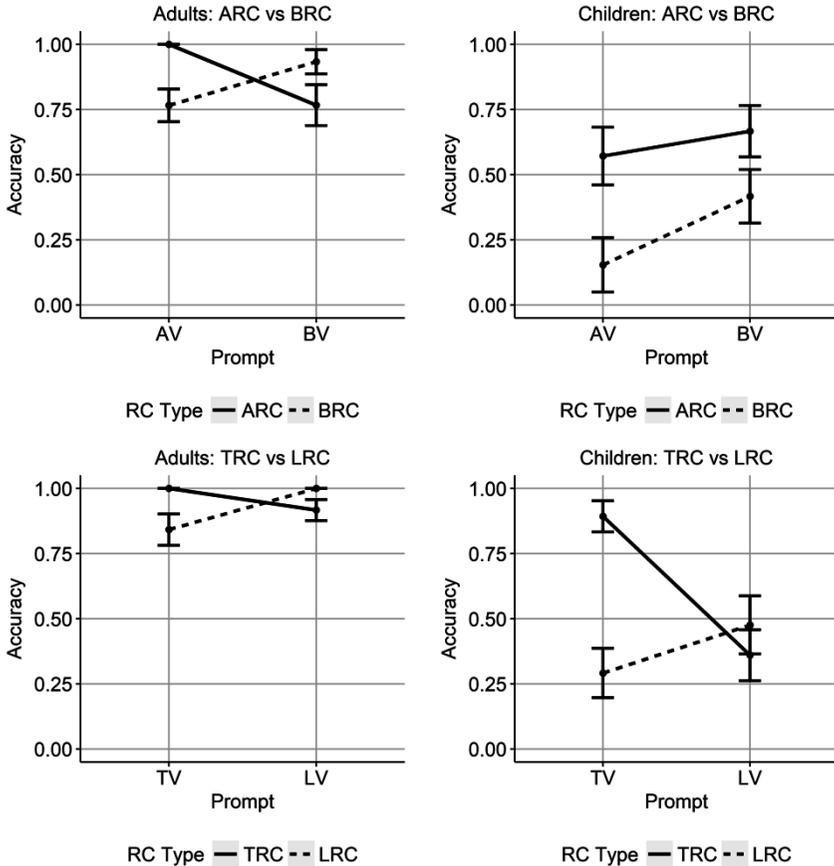
Figure 5. Graphs showing the interaction between RC Type and Prompt among adult ($n = 13$) and child ($n = 8$) participants on their grammatical RC responses. Error bars reflect standard errors.

Among adults, there was no RC type effect ($p = 0.95$), prompt effect ($p = 0.94$), or interaction between RC type and prompt ($p = 0.94$) for the ARC vs BRC comparison. For the TRC vs LRC comparison, there was no effect of RC type ($p = 0.44$) or prompt ($p = 0.48$), nor was there an interaction between RC type and prompt ($p = 0.99$)[4].

Results for the children revealed a significant effect of RC type ($\beta = -1.90$, $SD = 0.83$, $z = -2.30$, $p < 0.05$) for the ARC vs BRC comparison, but no effect of prompt ($p = 0.34$) and no interaction effect between the two predictors

---

[4] A *glm* model without random effects was used for the interaction of the RC type and prompt as the glmer models did not converge, possibly due to ceiling effects observed between the two conditions.

(p = 0.59). On the TRC vs LRC comparison, there was no effect of RC type (p = 0.32) or prompt (p = 0.47), nor was there an interaction effect (p = 0.30).

2.3. Experiment 2: Character Selection Task
2.3.1. Participants

Nineteen adults (age range 19–67, mean = 42.26) and eight children (age range 6;4–7;2, mean = 6;10) participated in the character selection task.

2.3.2. Method and materials

Participants were shown two pictures, and were asked to point to the character or item described by an audio-recorded relative clause pattern.

Participants were pseudo-randomly assigned to one of two lists. Each list contains a pseudo-randomized mix of the two pairs of relative clause types under study (12 items on ARC/BRC contrasts and 14 items on TRC/LRC contrasts) along with 10 other unrelated RCs. Each list also contained four practice items for transitive events that involve an animate and an inanimate element.
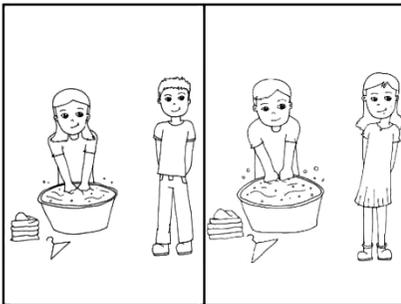

Figure 6. Sample ARC vs BRC comprehension item: *aŋ babaeŋ ipinaglaba naŋ lalaki* 'the girl for whom the boy did the wash'
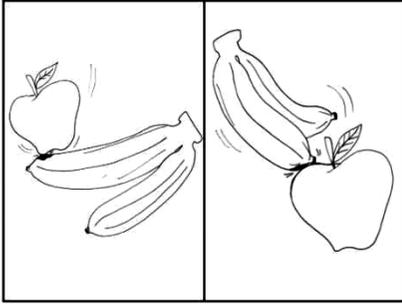
Figure 7. Sample TRC vs LRC comprehension item: *aŋ saging na nilaglagan naŋ mansanas* 'the banana on which (someone) dropped the apple'

Participant responses were categorized as *Target*, *Reversal Error*, *Head Error in the correct picture*, or *Head Error in the wrong picture*, as illustrated in Figures 8–9.
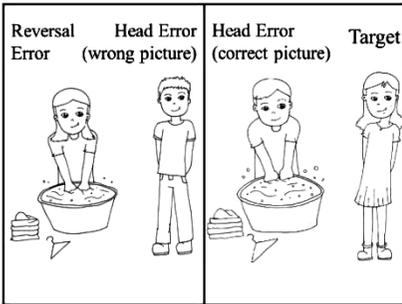

Figure 8. Response types for the ARC vs BRC comprehension item: *aŋ babaeŋ ipinaglaba naŋ lalaki* 'the girl for whom the boy did the wash'
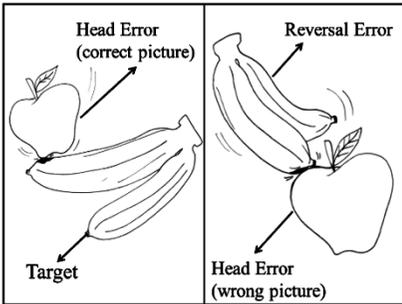

Figure 9. Response types for the TRC vs LRC comprehension item: *aŋ saging na nilaglagan naŋ mansanas* 'the banana on which (someone) dropped the apple'

2.3.3. Results

Asymmetries were found on the comprehension task (Tables 6–7). Both adults and children had higher accuracy rates on ARCs than BRCs and on TRCs than LRCs. Moreover, both groups exhibited high rates of reversal errors on the more difficult patterns, frequently interpreting BRCs as if they were ARCs and LRCs as TRCs.[5]

Table 6. Accuracy rates of adult participants in the character selection task ($n = 19$).

Response Category	Animate Heads		Inanimate Heads	
	ARC	BRC	TRC	LRC
Target	**95.61%**	**70.18%**	**92.48%**	**69.17%**
Reversal Errors	1.75%	18.42%	2.26%	14.29%
Head Errors				
Correct Picture	0.88%	8.77%	4.51%	10.53%
Wrong Picture	1.75%	2.63%	0.75%	6.02%

Table 7. Accuracy rates of child participants in the character selection task ($n = 8$).

Response Category	Animate Heads		Inanimate Heads	
	ARC	BRC	TRC	LRC
Target	**87.50%**	**18.75%**	**69.64%**	**32.14%**
Reversal Errors	0.00%	64.58%	5.36%	33.93%
Head Errors				
Correct Picture	4.17%	16.67%	17.86%	21.43%
Wrong Picture	8.33%	0.00%	7.14%	12.50%

The comprehension data were fitted to a mixed effects logistic regression model to determine differences between the two relative clause pairs under analysis. Using binary coding, target responses were coded with a "1", and the other types of responses with a "0". Similar procedures for setting up the models in the RC elicited production data were performed. The maximum model involved RC types as fixed effects, and participants and verbs as random effects. All random intercepts and slopes for all random effects were included. Contrast coding was applied. Terms were excluded one by one if the model did not converge, and a likelihood ratio test was performed for model comparison whenever necessary (Barr et al., 2013).

First, with the ARC vs BRC comparison, an effect of RC type was observed for both adults ($\beta = -3.18$, $SD = 0.77$, $z = -4.14$, $p < 0.001$) and children ($\beta = -4.27$, $SD = 0.86$, $z = -4.97$, $p < 0.001$). As for the TRC vs BRC

---

[5] The success rate of 70% by adults on certain RC patterns is in line with the results reported for English object RCs in studies of native English speakers; see, for example, DeDe (2015).

comparison, a similar effect of RC type was observed, for both adults ($\beta = 2.33$, $SD = 0.47$, $z = 4.95$, $p < 0.001$) and children ($\beta = 2.13$, $SD = 0.56$, $z = 3.78$, $p < 0.001$).

3. General Discussion

In sum, we found no differences on the two relative clause pairs among adults in the RC elicited production task. In contrast, the children showed a significant difference on the ARC vs BRC comparison, consistent with a more general tendency for success on ARCs, as reported by Tanaka (2016). Moreover, the absence of prompt effects and interactions in all four conditions suggest that priming was not a major factor in the children's performance.

The comprehension task yielded more clear-cut results, revealing asymmetries in each of the two relative clause pairs for both adults and children. In patterns with an animate head, both adults and children exhibited better performance on agent relative clauses compared to benefactive relative clauses. And in patterns with an inanimate head, they did better on theme relative clauses than on locative relative clauses.

It is not yet clear what factor, or combination of factors, is responsible for these asymmetries. However, we can at least rule out single-factor explanations involving structural prominence or morphological complexity, neither of which can account for the full range of developmental asymmetries in our study.

On a structural prominence account, the argument that is mostly highly placed in a syntactic structure is taken to be more accessible to relativization. On some accounts, agents occupy higher positions than locatives and benefactives, which in turn are higher than themes (Aldridge, 2004, 2012; Chen, 2016; Rackowski & Richards, 2005). These asymmetries align well with the superior success rate for ARCs compared to BRCs, but offer no insight into why TRCs yielded higher scores than LRCs.

The role of morphological complexity is also problematic in the absence of a plausible metric for use with sets of morphemes that differ from each other in more than one way. The situation for Tagalog is particularly challenging in this regard: the agent voice and the theme voice are both expressed by affixes that also indicate aspect; the locative voice is marked by a suffix that expresses only voice; and the benefactive voice is indicated by a bisyllabic prefix that can be further inflected by an infix that indicates aspect. The combination of positioning (prefix, infix or suffix), number of syllables (one or two), and the interaction with the expression of aspect effectively undermines attempts to explain the developmental facts we have reported by reference to any known complexity metric.

Yet another potential explanatory factor involves frequency, but here too the facts are less than straightforward. The theme voice occurs significantly more frequently than the agent voice in declarative clauses; yet Tanaka's (2016) study of ARCs and TRCs in which both arguments are animate showed a

significant advantage among children for ARCs in both production and comprehension.

Despite the limitations of our study due to small sample size and limited items, the findings of our research and the challenges associated with their explanation confirm the need to cast the net ever wider in the study of language acquisition. As is the case here, a syntactic construction (relative clauses) that has been the subject of a great deal of research in a small set of familiar European and Asian languages turns out to have unusual properties in a previously unstudied language, raising new questions for the study of language acquisition.

References

Aldridge, Edith. (2004). *Ergativity and word order in Austronesian languages* (Unpublished doctoral dissertation). Cornell University, Ithaca, NY.

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

Barr, Dale J., Levy, Roger, Scheepers, Christoph, & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Chen, Victoria. (2016). *"Pivot" ≠ absolute: Evidence from Formosan*. Proceedings of the 42nd Annual Meeting of the Berkeley Linguistic Society (BLS 42), Berkeley, California.

DeDe, Gayle. (2015). Effects of animacy on processing relative clauses in older and younger adults. *Quarterly Journal of Experimental Psychology*, *68*(3), 487-498.

Hsu, Chun-chieh Natalie, Hermon, Gabriella, & Zukowski, Andrea. (2009). Young children's production of head-final relative clauses: Elicited production data from Chinese children. *Journal of East Asian Linguistics*, *18*, 323–360.

Kim, Chae-Eun & O'Grady, William. (2015). Asymmetries in children's production of relative clauses: Data from English and Korean. *Journal of Child Language*, *43*, 1038–1071.

Kuznetsova, Alexandra, Brockhoff, Per Bruun, & Christensen, Rune Haubo Bojesen. (2017). *lmerTest: Tests in Linear Mixed Effects Models.* Version 2.0-36. https://CRAN.R-project.org/package=lmerTest

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rackowski, Andrea & Richards, Norvin. (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry*, *36*(4), 565-599.

Tanaka, Nozomi. (2016). *An asymmetry in the acquisition of Tagalog relative clauses* (Unpublished doctoral dissertation). University of Hawaiʻi, Honolulu, HI.

Proceedings of the 42nd annual Boston University Conference on Language Development

edited by Anne B. Bertolini
and Maxwell J. Kaplan

Cascadilla Press   Somerville, MA   2018