



Rowena Garcia* and Evan Kidd

Acquiring verb-argument structure in Tagalog: a multivariate corpus analysis of caregiver and child speech

<https://doi.org/10.1515/ling-2021-0107>

Received May 27, 2021; accepted November 13, 2021; published online July 7, 2022

Abstract: Western Austronesian languages have typologically rare but theoretically important voice systems that raise many questions about their learnability. While these languages have been featured prominently in the descriptive and typological literature, data on acquisition is sparse. In the current paper, we report on a variationist analysis of Tagalog child-directed speech using a newly collected corpus of caregiver-child interaction. We determined the constraints that condition voice use, voice selection, argument position, and thematic role assignment, thus providing the first quantitative analysis of verb argument structure variation in the language. We also examined whether children are sensitive to the constraints on variability. Our analyses showed that, despite the diversity of structures that children have to learn under Tagalog's voice system, there are unique factors that strongly predict the speakers' choice between the voice and word order alternations, with children's choices related to structure alternations being similar to what is available in their input. The results thus suggest that input distributions provide many cues to the acquisition of the Tagalog voice system, making it eminently learnable despite its apparent complexity.

Keywords: child-directed speech; corpus analysis; recursive partitioning; Tagalog; variation

1 Introduction

Verbs form the core of grammar, but as so-called 'hard' words (Gleitman et al. 2005), they constitute a complex domain for the child language learner to

*Corresponding author: Rowena Garcia, Language Development Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands, E-mail: rowena.garcia@mpi.nl. <https://orcid.org/0000-0003-1363-542X>

Evan Kidd, Language Development Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; Research School of Psychology, The Australian National University, Canberra, Australia; and ARC Centre of Excellence for the Dynamics of Language, Canberra, Australia, E-mail: evan.kidd@mpi.nl

negotiate. While children learning any language must determine how individual forms map onto sometimes subtle and unobservable meanings, individual languages also vary widely across different typological dimensions that influence how verbs and argument structure are expressed, meaning that children must rapidly identify language-specific cues in their input. In the current paper, we report on a corpus study of verb argument structure in Tagalog (Western Austronesian), a language that has a rare but typologically important *symmetrical voice* system, in which either agents or patients in transitive clauses can be made grammatically prominent without being marked (e.g., by passive morphology), and in which the order of arguments is relatively free.

We take a variationist approach to analyzing children's input and their productions, where we identify the variables that probabilistically condition the variation in voice-marking and word order in the input, and determine which of these variables children use in their own productions as they acquire the Tagalog voice system. The variationist method is well suited to an analysis of Tagalog verb argument structure, where the speaker is faced with multiple options concerning voice and word order choice. Although it is less often used in the study of child language acquisition (though see Budwig 1990; Labov 1989), the focus on explaining the determinants of variation in children's input and their early productions is thematically consistent with the assumption that children conduct distributional analyses over systematically organized input to acquire linguistic categories and structures (Saffran and Kirkham 2018; Tomasello 2003). Thus, identifying which variables condition variation in the input has the potential to reveal the cues which children use to acquire the target language.

1.1 The Tagalog voice system

Tagalog, one of the major languages of the Philippines, is a Malayo-Polynesian Austronesian language with over 25 million speakers (Eberhard et al. 2021). While there are alternative categorizations (e.g., Aldridge 2012; Chen and McDonnell 2019; Rackowski and Richards 2005; Schachter and Reid 2008), one dominant typological description of the language is that it constitutes a symmetrical voice language, such that the language has several basic transitive constructions where the verb is equally marked (Foley 1998; Himmelmann 2005a; Riesberg 2014). Accordingly, unlike in the active–passive alternation in languages like English, no argument is demoted to an oblique across the voice alternations (Riesberg and Primus 2015).

Consider a basic Tagalog sentence, which includes the predicate and the *ang*-phrase,¹ the sentence subject (Guilfoyle et al. 1992; Kroeger 1993a; but see Schachter 2015 for an alternative view). The predicate is commonly a verb that is marked for voice, aspect, and mood. Non-subject arguments as well as adjuncts are preceded by the marker *ng* (see Examples 1 to 4) or *sa*. Personal names are marked by *si*, *ni*, and *kay* instead of *ang*, *ng* and *sa*, respectively. Pronouns also appear in different *ang*, *ng*, and *sa* forms. A core feature of Tagalog is that the voice-marking on the verb assigns the thematic role of the *ang*-phrase (Himmelmann 2005b). In Example 1, the agent voice (AV) verbal infix *-um-* assigns the *ang*-phrase as the agent. In contrast, in Example 2, the patient voice (PV) verbal infix *-in-*² marks the *ang*-phrase as the patient (2). The assigned thematic roles are not influenced by the order of mention of the arguments, which means that agent voice in Examples (1) and (3) have the same meaning, as well as patient voice in Examples (2) and (4).

- (1) *K<um>i~kiliti ng bata ang babae*
 <AV>³IPFV~tickle NSBJ child SBJ woman
 ‘The woman is tickling a child.’
- (2) *K<in>i~kiliti ng bata ang babae*
 <PV>IPFV~tickle NSBJ child SBJ woman
 ‘The/A child is tickling the woman.’
- (3) *K<um>i~kiliti ang babae ng bata*
 <AV>IPFV~tickle SBJ woman NSBJ child
 ‘The woman is tickling a child.’
- (4) *K<in>i~kiliti ang babae ng bata*
 <PV>IPFV~tickle SBJ woman NSBJ child
 ‘The/A child is tickling the woman.’

Tagalog also has structures that are not verb-initial or where at least one argument occurs before the verb. Schachter and Otnes (1972) describe these as derived and minor structures. These include inversion or fronting constructions (i.e., the placement of an *ang*-phrase, adverb, or verb complement before a verb, through

¹ The *ang*-phrase has been also called pivot (LaPolla 2014), trigger (Schachter 2015), and topic (Schachter and Otnes 1972).

² Voice and mood are conflated in Tagalog.

³ AV refers to agent voice, PV to patient voice, SBJ to subject, NSBJ to non-subject, 1 to first person, 2 to second person, 3 to third person, SG to singular, PL to plural, INCL to inclusive, INF to infinitive, PFV to perfective, IPFV to imperfective, CONT to contemplated, LIN to linker, CAU to causative, EXIST to existential, and NEG to negation.

the use of *ay* constructions,⁴ or the linker *na* or *-ng*, or with or without the use of a pause between the first constituent and the verb), *wh*-questions, negations, modals (e.g., *gusto* ‘want’, *ayaw* ‘don’t want’), existential/non-existential words, relative clause constructions and cleft constructions (see Appendix A for examples).

Each language represents a unique challenge to the child language learner, but Tagalog verb argument structure appears to be a particularly complex puzzle. In the remainder of the introduction, we review the range of variables that influence voice selection and word order in Tagalog. Inspired by variationist approaches to grammar (e.g., Bresnan 2007; Szmrecsanyi 2017; Tagliamonte and Baayen 2012; Weiner and Labov 1983), we identify a set of variables that likely influence voice selection and argument ordering, under the assumption that such variables restrict the hypothesis space and act as cues for children to acquire the verb argument structure of Tagalog. We then test whether these conditioning factors are available in the input and if they govern children’s own productions, using data from a newly collected corpus of Tagalog caregiver-child interactions (Garcia 2021). In doing so, we select variables by drawing heavily from the general linguistics literature on Tagalog grammar, since no comprehensive corpus studies of verb argument structure in Tagalog exist. Thus, while we are primarily concerned with how children’s input is structured and might make Tagalog verb argument structure learnable, our results are broadly applicable to the linguistics of Tagalog in general. With this in mind, we next review four domains that bear upon our analyses and which vary in the language: (i) voice-marking on the verb (versus non-use), (ii) selection of voice type, (iii) word order, and (iv) thematic role assignment.

1.1.1 Verbal inflection and voice

1.1.1.1 Voice inflection

Although using a verb that is not marked for voice is ungrammatical in Tagalog (aside from recent perfectives), there is evidence that the input that children receive also contains verbs that are not inflected for voice, which may be a prominent feature of Tagalog child-directed speech. For instance, Garcia et al. (2019) found that a quarter of child-directed utterances with highly causative transitive verbs were not inflected for voice. In another study, Garcia et al. (2021) showed that approximately half of uninflected highly causative verbs in child-directed speech were imperative or hortative utterances. Assuming that child-directed speech reflects the caregiver’s adjustments to the child’s current language

⁴ *Ang babae ay kumikiliti ng bata* is the *ay*-inverted structure for Examples 1 and 3 that mean, ‘The woman is tickling a child.’

level (Phillips 1973; Pine 1994), and given that Tagalog-speaking children initially produce mostly uninflected verbs (Marzan 2013), it could be the case that caregivers simplify the input by using root words instead, and increase the use of voice-inflected verbs only when the child already uses more grammatical morphemes.

1.1.1.2 Voice type

One clear source of complexity in Tagalog verb acquisition comes from verb inflection: the language contains a complex system of affixes (prefixes, suffixes, and infixes), which mark voice, aspect, and mood. In the agent voice (the *ang*-phrase is an agent), the verb can be marked by the infix *-um-*, or the prefixes *mag-* and *maN⁵-*. In the patient voice (Himmelman's [2005b] *undergoer voices*, where the *ang*-phrase is not an agent, but is usually a patient, recipient, instrument or beneficiary), the verb is marked by the suffix *-in*, the suffix *-an*, or the prefix *i-*. There are also corresponding agent (*maka-*) and patient voice (*ma-*, *ka-*) affixes for stative and potentive verbs, which depict events without a volitional or controlling agent (see Himmelman 2005b; Appendix B for an overview of the affixes). Analyses of a written corpus (Cooreman et al. 1984) and child-directed speech (Garcia et al. 2019) show that in transitive sentences, the patient voice is more frequent than the agent voice.

Within each voice, speakers also have to identify which specific verbal marker to use. The choice of the agent voice marker seems to depend on the base form of the verb but sometimes indicates semantic differences (Himmelman 2005b; Latrouite 2001). For example, *mag-* and *maN-* indicate a greater frequency or intensity compared to *-um-*. For the patient voice, marker choice has been claimed to be dependent on how (directly) affected the non-agent argument is (Himmelman 1987), such that *-in* is used when the non-agent is directly affected or more affected, and *-an* and *i-* when the non-agent is indirectly or less affected. The choice of affix within each voice adds another dimension to what children have to learn, but in this paper we focus only on their acquisition of the voice alternation.

According to Himmelman (2005b), despite the high number of possible inflections, a particular verb is usually used with only one or two voice affix/es. However, the factors that determine the final choice of voice is still poorly understood. One robust rule involves definiteness: *ang*-phrases are always interpreted as definite, therefore the patient voice is used in simple main clause constructions with definite patients or undergoers/non-agent arguments (Himmelman 2005b; Latrouite 2011; others refer to specificity of the non-agent argument; Adams and Manaster-Ramer 1988; Maclachlan and Nakamura 1997; Rackowski and Richards 2005). However, the rule involving definiteness no longer

5 *N* refers to a nasal phoneme.

applies to relative clauses, and existential and cleft constructions (Himmelman 2005b). Himmelman claims that if the verb is not in the predicate position (non-verb-initial), different rules obtain and voice is determined by the syntax of the construction instead.

Others have also claimed that the patient voice is preferred for coding events that are highly semantically transitive (causative), based on Hopper and Thompson's (1980) criteria (e.g., volitionality of the agent and affectedness of the non-agent) (Katagiri 2005; Latrouite 2011; Nolasco 2005). This hypothesis is supported by written corpus data (Cooreman et al. 1984) and spontaneous child-directed speech (Garcia et al. 2019). Additionally, it has been claimed that animacy of the argument affects voice choice, with the patient voice being strongly favored when there is an animate non-agent (Himmelman 2005b; Latrouite 2011). Latrouite further argues that the agent voice is not acceptable given a human non-agent. Experimentally, Tanaka et al. (2014) showed that adult native speakers have a stronger preference for the patient voice when the non-agent is definite and animate than when it is indefinite and inanimate.

1.1.2 Word order and thematic role assignment

1.1.2.1 Word order

Tagalog's canonical order is predicate-initial, but the order of the arguments is relatively free (Schachter 2015). For verb-initial sentences with two arguments, there have been claims that Tagalog's canonical order is verb-*ang-ng* (Aldridge 2002), verb-*ng-ang* (Billings 2005), or both verb-*ang-ng* and verb-*ng-ang* for the agent voice but only verb-*ng-ang* for the patient voice (Guilfoyle et al. 1992; Kroeger 1993b). Others have argued that the canonical word order is agent-initial (Buenaventura Naylor 1975; Manuelli 2010; Schachter 2015). Thus, the hypothesis space is large and we note that there have been very few corpus studies of spoken Tagalog that have contributed to this debate.

In sentences with non-pronominal arguments, Kroeger (1993b) proposed three principles guiding the order of the arguments: (1) the agent tends to be the first argument, (2) the *ang*-phrase tends to be the last, and (3) the heavier argument (longer constituent) follows the lighter argument (shorter constituent). In sentences with one pronominal argument, the pronoun is mentioned immediately after the verb (Billings 2005; Schachter and Otanes 1972). Given two pronouns, the order of mention depends on syllable-length, with monosyllabic pronouns preceding disyllabic pronouns (Reid and Liao 2004; Riesberg et al. 2019).

While language-specific variables no doubt influence word order, other conceptual variables independent of Tagalog grammar may also play a role. Accordingly, crosslinguistic studies have shown that animates tend to precede

inanimates (Bock et al. 1992; Ferreira 1994; Kempen and Harbusch 2004; Minkoff 2000; Tanaka et al. 2011; Van Valin and LaPolla 1997), either via the tendency to place animate arguments in prominent syntactic positions (in S-initial languages) or because speakers prefer to produce them first as they are conceptually accessible. Additionally, Riesberg et al. (2019) have proposed that even in languages with a flexible order of arguments in basic transitive constructions, such as in symmetrical voice languages, there is still a processing bias to position the (animate) agent as the first argument. Non-agent-first constructions supposedly exist due to other ordering principles, which could be in competition or in conflict with this agent-first preference.

1.1.2.2 Thematic roles

The symmetrical voice system in Tagalog means that children must learn that the function of a noun marker can change depending on voice. That is, *ang* in the agent voice marks an agent, but in the patient voice, marks a non-agent. There are claims that, crosslinguistically, thematic roles are influenced by animacy and referentiality or (pro)nominality (Croft 2003; Silverstein 1976). Animate entities prototypically act as agents while inanimate entities are more likely to act as non-agents (Dowty 1991; Hopper and Thompson 1980). An animate entity's ability to volitionally instigate actions make them ideal agents, and the lack of this ability in inanimate entities make them less ideal. In terms of referentiality, pronouns are more likely to code for agents in comparison to nouns (Croft 2003). In Tagalog child-directed speech, it has been shown that in sentences with one pronoun and one full noun phrase, the pronoun coded the agent 97% of the time (Garcia et al. 2018).

In Tagalog transitive sentences with only one argument (the other argument elided), there is evidence that this sole argument is most likely the agent. In child-directed speech, Garcia et al. (2019) showed that there is a general agent-initial preference, although this seemed to be influenced by voice: transitive utterances with agent voice verbs were 95% agent-initial or agent-only, while those with patient voice verbs were 85% agent-initial or agent-only. However, utterances with pronominal and nominal arguments were mixed in the counts; and only verb-initial utterances with highly causative verbs were included.

1.2 Acquisition of the Tagalog voice system

There are currently only a few studies on children's acquisition of the Tagalog voice system, and thus many of the variables we analyze in the current paper have

never before been investigated, either in child-directed speech or in children's productions. We briefly review this work here.

Based on spontaneous speech samples from five Tagalog-English bilingual children (1;2–3;7; Marzan 2013), Marzan and colleagues (2014) reported that 69% of children's verbs were not inflected for voice nor aspect.⁶ Based on data from four of the children in Marzan's corpus, Garcia (2016) reported that 59% of voice-marked verbs (including both intransitives and transitives) produced by children were in the patient voice, while 41% were in the agent voice.

Children's competency in both voice and noun-marking improves with age. Through a picture description and sentence completion task using *ay*-inversion sentences (e.g., *Ang bata ay ...* "The child is ..."), Segalowitz and Galang (1978) found that 3-year-old Tagalog-speaking children were less accurate in marking voice on the verb compared to older children (5- and 7-year-olds). However, the proportion of correct voice-marking was similar for agent and patient voice. Another study using a similar task also showed that 3-year-olds were less accurate in voice-marking compared to older children (based on mean scores), but, in contrast to the Segalowitz and Galang study, all groups of children (3-, 5-, 7-, 8-year-olds) were more accurate in marking the patient voice on the verb compared to the agent voice (Galang 1982). Moreover, in picture description and sentence completion tasks using verb-initial sentences, 3- to 7-year-old Tagalog-speaking children made more errors in noun-marking given verbal prompts in the agent voice compared to prompts in the patient voice (Garcia et al. 2018; Garcia and Kidd 2020).

Focusing on word order, Garcia (2016) also reported that 83% of children's agent voice utterances in Marzan's (2013) corpus were agent-only, 14% were agent-before-*non-agent*, and the remaining 3% were *non-agent*-only or *non-agent*-initial. Patient voice utterances were more likely to contain two overt arguments compared to agent voice utterances. However, similar to the agent voice, most of the utterances were agent-initial: 49% agent-only, 45% agent-before-*non-agent*, and the remaining 6% were *non-agent*-only or *non-agent*-initial. Using a picture description task, Bautista (1983) found that 2- to 4-year-old children were more likely to produce an agent-initial order (88%) than a patient-initial order (12%); though the interaction of word order and voice was not reported. For older children (mean age = 5;5 years), Tanaka (2016) found that the most preferred structure for describing transitive events was the patient voice agent-initial structure (verb-*ng-ang*). Children also showed an agent-initial preference in the agent voice.

⁶ However, this number also includes children's verbal productions in English.

1.3 The current study

The review of the literature shows that, despite the diversity of structures that children have to learn under Tagalog's voice system, there are identifiable factors that likely influence speakers' choice between the voice, word order, and thematic role alternations. We currently lack a good understanding of how these factors influence variability in verb argument structure in Tagalog, as most of the claims on Tagalog voice selection and word order have not yet been quantitatively assessed. Moreover, it is also not clear how this variability influences the acquisition of the system. Our working assumption is that, despite the likely variability in the system, there is significant *systematicity* such that those variables that condition variation in the input act as cues to children in acquisition. As such, in the current study, we report a variationist analysis of corpus data collected from observational recordings of 20 caregiver-child dyads, where the children (aged 2–4 years) were acquiring Tagalog as a first language. We aimed to (i) determine which factors condition variation and are thus available as cues in the input to children, and (ii) identify which cues children use as starting points for acquiring Tagalog's voice system.

2 Method

2.1 Participants

Twenty Tagalog-speaking caregiver-child dyads from Caloocan city participated in the study. The children's ages ranged from 2;0 to 4;0 years. The caregivers were mostly the mothers of the children (12). The rest were grandmothers (5), fathers (2), and one grandfather. The mean age of the caregivers was 36 years (range: 21–72 years). Participants were selected because Tagalog was the dominant language in their households. Moreover, the caregivers were all native Tagalog speakers; however, they also reported proficiency in English and/or other Philippine-languages, which is typical for Tagalog speakers from Metro Manila (Amora et al. 2020; Mahboob and Cruz 2013; Philippine Statistics Authority 2003). Some caregivers reported proficiency in Kapampangan (3), Bikol (2), Cebuano (1), Ilocano (1), and Pangasinan (1). Their educational attainment was mixed: university (3), some university education (8), high school (7), some high school (1), and elementary (1).

2.2 Materials

In order to elicit conversations, the caregiver-child pairs were provided with toys and books. The toys included dolls, animal figures, kitchen set, doctor set, blocks, cars, furniture miniatures, and a magic slate. The set also included the wordless picture storybook *Frog, Where are You?* (Mayer 1969). Additionally, we compiled different pictures of causative and non-causative actions with varying animacy of the entities, in the hope that they would elicit a greater amount of basic clauses. A video camera with a microphone was used to record the sessions.

2.3 Procedure

The first author (Tagalog native speaker) encouraged the caregivers to use the materials provided to engage with the children. No other instruction was given. The researcher did not interact with the dyad unless it was completely necessary. Each recording session lasted 60 min.

The recordings took place in the families' living rooms (which for some was the same as the bedroom). Caregivers were informed beforehand that there should be no one else in the room aside from them, the children, and the researcher. Given that extended family members typically live together in the Philippines (Chen et al. 2017), and a mean household size of 4.4 persons (Philippine Statistics Authority 2016), one-on-one interactions between children and caregivers might not be as common compared to countries with a smaller average household size. While we acknowledge that the recording sessions do not closely mimic the exact experience of the Tagalog-learning child, our method fulfilled our aim of eliciting substantial amounts of child-directed speech in a semi-naturalistic setting, with the minimal number of participants easing transcription.

2.4 Data annotation

Two research assistants (Linguistics graduate and Speech Pathology graduate who are native speakers of Tagalog) transcribed the video recordings using ELAN (version 5.9, 2020). The transcription rules were adapted from the minCHAT format (MacWhinney 2000) and were similar to the DARCLE Annotation Scheme (Casillas et al. 2017). Verb phrases were annotated in ELAN by one of the research assistants (and rechecked by the first author) for type of utterance (e.g., declarative, imperative), sentence constituent (e.g., verb, negation word, existential word, etc.), verb inflection (voice, aspect, and mood inflection), causativity of the verb, thematic

role of the arguments, the noun-marking used, as well as the nominality and animacy of the nouns. Verb phrases without any Tagalog word, e.g., “Look! Look!” were not further annotated.

The verbs were further categorized into intransitive, transitive non-causative, and transitive causative. Ditransitives were marked as transitives. We judged causativity based on Hopper and Thompson’s (1980) criteria such as the volitionality of the agent and the affectedness of the non-agent argument (see Appendix C for the complete list). For example, *tulak* ‘push’ was considered causative but not *hanap* ‘look for.’ A few verb bases like *luto* ‘cook’, *laro* ‘play’, *inom* ‘drink’, and *kain* ‘eat’ were categorized as either intransitive or transitive causative. Verb bases like *sing* ‘kanta’ and *basa* ‘read’ were categorized as either intransitive or transitive non-causative. Agent voice-inflected verbs without a non-agent argument were considered intransitives (5a). Agent voice-inflected verbs with a non-agent argument were considered transitives (5b). Utterances with patient voice verbs were considered transitives, regardless of whether the non-agent (*ang*-phrase) was dropped (5c) or not (5d). Non-voice-inflected verbs with a *ng*-form agent were considered transitives (5e); and those with an *ang*-form agent but no non-agent argument were considered intransitives (5f), as well as those which did not have any argument (5g). Our intransitives are similar to Schachter and Otnes’ (1972) *simple intransitives*, although we considered what they call *pseudo-transitive* (verbs that accept an agent adjunct), *adjunctive* (verbs that accept adjuncts except for agent adjuncts), and *adjunctive pseudo-transitive* (verbs that accept both agent adjunct and other adjunct type) as transitives. Additionally, we categorized agent voice-marked verbs with non-agent arguments as transitives, which Nolasco and Saclot (2005) consider as grammatically intransitive but semantically transitive verbs.

- (5) a. *K<um>ain sila.*
 <AV>eat.PFV 3PL.SUBJ
 ‘They ate.’ (C16)⁷
- b. *K<um>a~kain ka ng gulay?*
 <AV>IPFV~eat 2SG.SUBJ NSBJ vegetable/s
 ‘Do you eat vegetables?’ (C2)
- c. *Dali, kain-in mo na.*
 quickly eat-PV.INF 2SG.NSBJ already
 ‘Quickly, you eat (it) already’ (C3)
- d. *Kain-in mo na raw yung egg.*
 eat-PV.INF 2SG.NSBJ already SBJ egg
 ‘You eat the egg already’ (C3)

⁷ C stands for ‘caregiver’ and the number refers to the participant ID.

- e. *Kain mo.*
eat 2SG.NSBJ
'You eat (it)' (C1)
- f. *Kain na tayo.*
eat already 1PL.INCL.SBJ
'Let's eat already' (C8)
- g. *Kain na muna.*
eat already first
'Eat first.' (C8)

The verb base *gawa* 'do/create/fix' was classified as transitive causative (6a) or transitive non-causative (6b; more frequent occurrence) depending on the context.

- (6) a. *G<in>a~gawa 'yung kalsada.*
<PV>IPFV~fix SUBJ road
'The road is being fixed.' (C16)
- b. *Ano-ng g<in>a~gawa ni Ate?*
what-LIN <PV>IPFV~do NSBJ big sister
'What is big sister doing (instead of making)?' (C19)

The frequent verb base *sabi* 'say' was considered as transitive non-causative (7) and verbs *nangyari/nangyayari* 'happened/happening' were considered as intransitive (8).

- (7) *Tapos, sabih-in mo daw "palaka".*
then say-PV.INF 2.SG.NSBJ frog
'Then, you say "frog"' (C20)
- (8) *Ano-ng nangyari sa kabayo?*
what-LIN happen.PFV NSBJ horse
'What happened to the horse?' (C5)

In terms of thematic roles, we focused on core arguments: agent and non-agents. Agent and non-agent in the annotations followed Dowty's (1991) more encompassing proto-agent and proto-patient roles, respectively. A proto-agent has the following properties: exists independently of the event or the action, volitionally involved, can perceive or feel things, causes a change in an event or in a state of another entity, or moves in relation to another entity. Contrary to the proto-agent, a proto-patient is causally affected by another entity, stationary in relation to another entity, involved in telic events (e.g., She walked *the park*), undergoes a change of state, and does not exist outside of the event (e.g., The girl drew *a zebra*). For example, the sole argument of the intransitive verb *nabali* 'got broken' was

marked as a non-agent argument (9, undergoes a change of state). Adjuncts were not further annotated. Existentials and non-existentials without complements (e.g., *may* ‘there is’ instead of *may bata* ‘there is a child’), as well as the *ang*-phrase in headless relative clauses, were considered as arguments. Question words *sino* ‘who?’ and *ano* ‘what?’ were also considered as arguments. Given the pre-verbal position of question words, this means that the high frequency of *ano* questions in the patient voice (refers to the non-agent argument) may result in a higher number of non-agent-initial sentences compared to previous analyses which included only verb-initial declarative sentences (Garcia et al. 2019; Garcia et al. 2021). Additionally, considering question words as arguments may affect the position of monosyllabic pronouns (i.e., as the question word will be the first argument instead). Finally, since we considered question words as *ang*-phrases, this may result in a higher frequency of *ang*-first sentences compared to counts limited to verb-initial sentences.

- (9) *Na-bali* *’yung puno.*
 <PV>break.PFV SUBJ tree
 ‘The tree got broken.’ (C12)

Arguments were marked either as agent or non-agent. We did not focus on determining whether an argument was a patient or a beneficiary, so the first argument in Example 10 was simply considered as a non-agent. Additionally, we considered the referent of *anong* ‘what + linker’ in Example 11 as a non-agent.

- (10) *Tulong-an* *ka* *na* *ni* *Mama?*
 help-PV.INF 2.SG.SUBJ already NSBJ Mama
 ‘(Should) Mama already help you?’ (C6)
- (11) *Ano-ng* *g<in>a~gawa* *ng* *bird?*
 what-LIN <PV>IPFV~do NSBJ bird
 ‘What is the bird doing?’ (C3)

We also marked whether the argument was a pronoun, a noun, a clause, or the referent of a question word or a relative clause. Additionally, the specific pronouns and questions words used, as well as the markers used for the nouns and clauses were annotated. Regarding animacy, the arguments were classified into animate, inanimate, and activity or referent (such as “frog” in Example 7 and “ano-ng” in Example 11).

2.5 Data analysis

The ELAN annotations were exported to txt files, and the pre-processing, calculations and statistical analyses were performed in R statistical software (version 4.1.1, R Core Team 2016).

2.5.1 Child-directed and children's speech

Utterances were first categorized into verb-initial and non-verb-initial, using R. By non-verb-initial, we refer to utterances which have an argument that occurs before the verb (12a, in comparison to 12b which is verb-initial). We also identified utterances with animate non-agents and definite non-agents (e.g., pronominal non-agents, non-agents marked by *ang*, *sa*, *ni*, *kay*).

- (12) a. *Hindi niya po h<in>a~habol.*
 NEG 3.SG.NSBJ <PV>IPFV~chase
 'He/She is not chasing (it)' (C12)
- b. *H<in>abol ni Nanay.*
 <PV>chase.PFV NSBJ Mother
 'Mother chased (it).' (C12)

Child-directed productions were analyzed separately from children's speech. The children's mean length of utterance (MLU) based on morphemes was calculated from 100 intelligible utterances from the 30th minute mark of the recording (or from the whole recording if need be; see Table 1; note that participants 8 and 13 had only 88 and 85 intelligible utterances respectively). Notably, utterances without verbs were also included in the MLU calculation but not included in the

Table 1: Participants' age and mean length of utterance (MLU) based on morphemes on 100 utterances. Participants with an MLU below 2 are in Group 1, and those with an MLU above 2 are in Group 2.

Group 1			Group 2		
Participant	MLU	Age	Participant	MLU	Age
3	1.04	2;2	20	2.07	3;2
8	1.10	2;1	2	2.14	3;10
4	1.24	3;4	15	2.66	3;2
9	1.24	2;10	14	2.89	2;8
1	1.31	2;9	7	2.92	3;7
13	1.38	2;3	5	2.94	3;7
18	1.44	3;3	19	3.09	3;3
17	1.47	2;2	16	3.56	3;10
10	1.58	3;10			
12	1.59	2;7			
11	1.65	2;9			
6	1.69	2;1			
Average	1.39	2;8	Average	2.78	3;4

subsequent analyses. Immediate repetitions of caregiver utterances, the singing of memorized songs, and sentences which did not contain Tagalog words (e.g., “What’s this?”) were excluded from MLU counts and all analyses. MLU was calculated by exporting the utterances first to CLAN (MacWhinney 2000), then creating a morphological tier using a manually-annotated MOR lexicon file (containing all the unique words in the 100 intelligible utterances from each child), and then using CLAN’s MLU command.

Children were divided into 2 groups based on MLU. Those with an MLU of less than 2 morphemes were placed in Group 1 (12 children), and those with an MLU greater than 2 morphemes were in Group 2 (8 children). Dividing the children into developmental stages based on MLU is common in child language studies (Brown 1973), since the considerable individual differences in children’s rate of acquisition means that age-based groups are a comparatively poor proxy for developmental level (Kidd and Donnelly 2020). While any cut-off would be arbitrary, setting it at MLU = 2 likely marks a significant developmental step in a language like Tagalog, where the grammar requires morphological marking on both the verb and noun phrases (NPs).

2.5.2 Conditional inference trees and conditional random forests

We performed non-parametric analyses, namely conditional inference trees and conditional random forests (Strobl et al. 2009), to determine the factors that influence the selection of voice inflection (voice-marked or not), voice type, argument position, and thematic role of first argument (NP1) in the child-directed speech and in children’s speech. Conditional inference trees and conditional random forests are useful for identifying the variables in corpus data that are associated with the choice between linguistic variants, with the former being more useful for explanation and interpretation, and the latter for prediction (Levshina 2021). These tree-based methods involve repeatedly partitioning the data into two branches such that the split leads to the best increase of classification accuracy in predicting the dependent variable (Gries 2020). It is considered to be less affected by sparsity of data and collinearity compared to logistic regression (Tagliamonte and Baayen 2012). Additionally, the tree-based visualization is easier to interpret compared to coefficient tables from logistic regression models (Baayen et al. 2013).

From the conditional random forests, we obtained conditional variable importance scores. These scores show how important a variable or factor is in predicting the outcome (e.g., selection of voice type), by determining how much the prediction deteriorates from randomly reshuffling a factor’s values (Levshina 2016). A strong association between the variable and the outcome leads to a strong negative effect of the reshuffling. The importance score is the mean of the scores of

all trees in the conditional random forest. Importantly, it is called *conditional* as the influence of each factor is computed conditionally on the other factors. Unimportant factors have scores that are close to zero.

We used the party package (version 1.3-8, Hothorn et al. 2006) to fit the conditional inference trees and conditional random forests. For the conditional inference tree, the default settings of `ctree_control` were taken, except for `mincriterion` (0.99) and `maxdepth` (3) for the child-directed speech, and `mincriterion` (0.99) for children's speech. For the conditional random forest, the default settings of `cforest_unbiased` were used, except for `mtry` (2) and `n tree` (1,000).⁸ We used the `pdp` package (version 0.7.0) to create partial dependence plots to determine the influence of individual variables on the dependent variable. To test for the models' fit, we calculated the tree's classification accuracy (number of correct prediction divided by the number of observations), and we used the `Hmisc` package (version 4.5-0) to obtain the concordance index (*C-index*). The *C-index* refers to the proportion of randomly sampled observations with outcome A which the model predicts to have a higher probability of having outcome A compared to a randomly sampled observations with outcome B; a value of 0.5 does not discriminate between the outcomes while a value of 1 shows perfect discrimination (Levshina 2021). For calculating the classification accuracy and *C-index* of the conditional random forests, we used the samples which were not used during subsampling (i.e., out-of-bag samples). The data and the R scripts for the analyses can be found at <https://osf.io/q568p/> (DOI 10.17605/OSF.IO/Q568P).

3 Results and discussion

There were a total of 9,401 verb phrases in the child-directed speech. Overall, 15% of verb phrases did not contain a single argument. The category of transitive causative verbs was the most frequent, (39%), followed by intransitive verbs (33%), and then by transitive non-causative verbs (28%). Verb-initial utterances comprised 80% of the data, while the remaining 20% were non-verb-initial. Of the non-verb-initial utterances, 52% contained questions, 17% involved fronting an argument (but none had *ay*-inversion), 11% contained modal verbs, 8% contained negations, and 4% had an existential/non-existential word. Another 4% were marked as clefts, and 4% as relative clauses.

In children's speech, there were a total of 1,453 verb phrases. Thirty-four percent (487 utterances) were produced by children from Group 1 (i.e., MLU below 2), and the remaining 66% (966 utterances) were from Group 2 (i.e., MLU higher

⁸ Using `mtry = 3` yielded similar results.

than 2). In total, 63% of verb phrases produced by Group 1 did not contain a single argument, while in Group 2 the number was a considerably lower 29%. In Group 1, most of the utterances were intransitive (56%), followed by transitive causatives (39%), and transitive non-causatives (5%). In Group 2, there was almost the same number of intransitives (43%) and transitive causatives (41%). The remaining 16% were transitive non-causatives. In Group 1, 98% of the utterances were verb-initial; they produced only 12 non-verb-initial utterances. In Group 2, 89% were verb-initial. Of Group 2's 110 non-verb-initial utterances, 37% involved fronting an argument (but none had *ay*-inversion), 18% contained negations, 16% contained questions, 14% contained modal verbs, and 12% had an existential/non-existential word. Another 3% were marked as relative clauses.

3.1 Verbal inflection and voice

3.1.1 Voice inflection

This analysis investigated the factors that determine whether or not a verb is inflected for voice. We tested the following factors: utterance type (verb-initial or not), sentence type (imperative or not), MLU group of the child (Group 1, Group 2), and the individual variation among the speakers (participant ID). All utterances with verbs were included. Following the past literature, we predicted that verbs are likely to be unmarked for voice in imperative constructions (Garcia et al. 2021), in simpler utterances (canonical verb-initial vs. non-verb-initial), and in the speech of caregivers to children whose productions are simpler and shorter (Phillips 1973; Pine 1994).

3.1.1.1 Child-directed speech

Out of the 9,401 utterances with verbs, 35% had verbs which were not inflected for voice. The conditional inference tree on utterances with verbs (Figure 1) shows that the first split and two others are based on participant ID (Nodes 1, 4, 7), with those on the left branch producing a similar number of voice-marked and non-voice-marked verbs in general, while those on the right branch produce mostly voice-marked verbs. Those on the left branch (Node 2) and five from the right branch (Node 11) show a stronger tendency to mark voice in non-verb-initial utterances, in comparison to verb-initial utterances. Additionally, six of those on the right branch (Node 8) show a higher tendency to mark voice in non-imperative compared to imperative utterances. Groups based on children's MLU do not produce a significant split in the first three split levels.

The conditional variable importance scores from the conditional random forest model (Figure 2) shows that participant ID is the most important predictor for

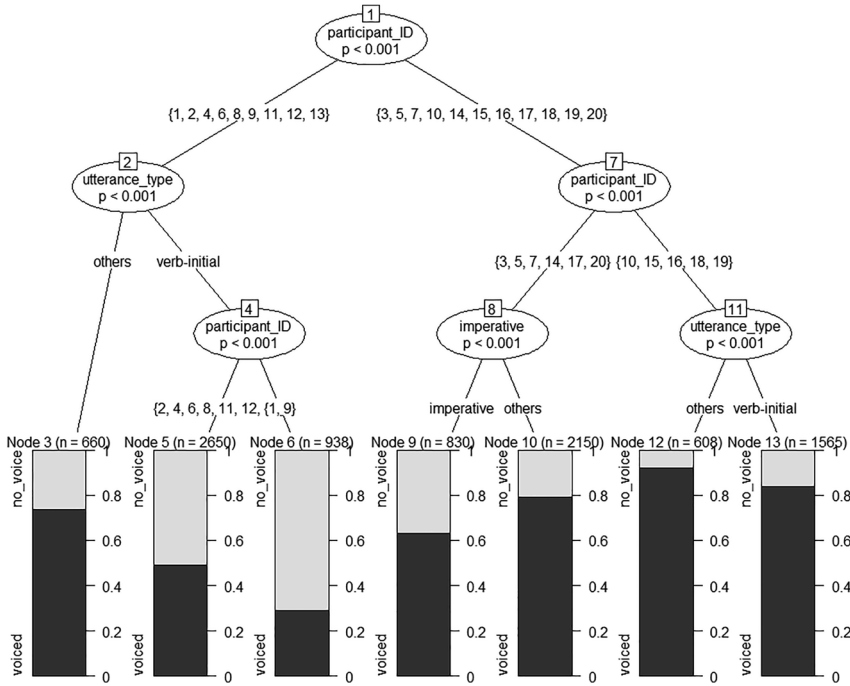


Figure 1: Conditional inference tree for voice inflection in child-directed speech. Numbers on top of the circles indicate the node numbers. The p -values show the significance of the association between the independent variable in the circle and voice inflection. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, n indicates the number of observations, and the bar graphs indicate the fraction of voice-marked verbs.

voice inflection, similar to the first split in the conditional inference tree. This is followed by whether or not the utterance is imperative, or verb-initial. Children's group (based on MLU) is not significant.

The models have satisfactory predictive power. The conditional inference tree has a classification accuracy of 0.70, while the C -index is 0.73. The conditional random forest's out-of-bag classification accuracy is 0.72, and the out-of-bag C -index is 0.75.

These results support two of our predictions. We predicted that caregivers would simplify the input and adjust to the child's language level (Phillips 1973; Pine 1994). Specifically, we predicted that if the use of root words is related to simplifying the input, there would be a higher tendency to use the base forms of verbs in canonical verb-initial sentences compared to more complex non-verb-initial sentences, and this is what we found. Verbs are less likely to be marked

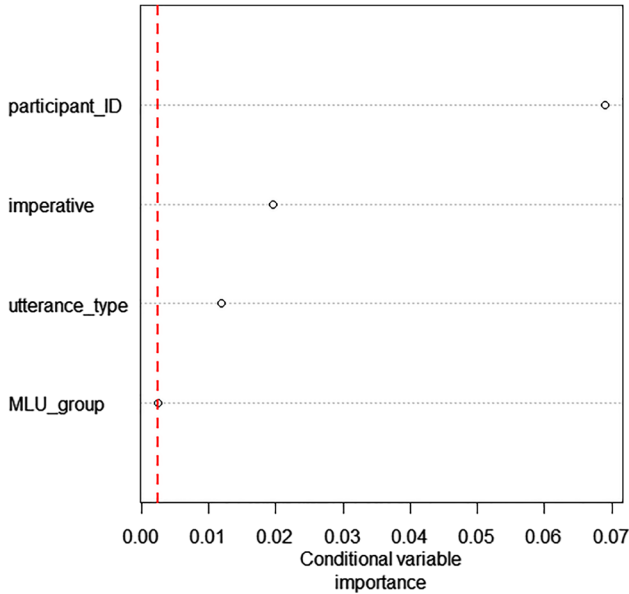


Figure 2: Conditional variable importance scores of the conditional random forest for voice inflection in child-directed speech. Predictors with dots to the right of the red vertical line are significant.

for voice in simpler utterances (canonical verb-initial vs. non-verb-initial). Additionally, similar to the finding of Garcia et al. (2021), we have evidence that root words are used more often for imperative or hortative utterances than for non-imperative utterances. However, we do not have evidence that caregivers increase the use of voice-inflected verbs as the children’s productive proficiency improves.⁹ Instead, we found that individual speaker variation is the strongest predictor of inflecting voice on the verbs. Therefore, the availability of voice-marking in the input that the child receives is mostly dependent on the caregiver, which may be suggestive of individual variability in the degree to which caregivers hold folk theories of language acquisition, thus leading to differential investment in the Tagalog child-directed speech register. This interesting finding awaits further research.

3.1.1.2 Children’s speech

Out of Group 1’s 487 verbal utterances, 81% had verbs which were not inflected for voice. From Group 2’s utterances, only 35% were not inflected for voice. Thus, children are more likely to mark voice as they become more proficient.

⁹ As pointed out by an anonymous reviewer, it is also possible that caregivers attune to the child’s proficiency, but our MLU measure captures only part of this proficiency.

The conditional inference tree on utterances with verbs (Figure 3) shows that the first and two other splits (Nodes 1, 2, 10) are based on participant ID. Those on the left branch produce more voice-marked verbs than non-voice-marked verbs, while those on the right produce mostly non-voice-marked verbs. The participants on the left branch (Nodes 3, 6) also show a higher tendency to mark voice on non-imperative utterances compared to imperatives. MLU group also makes a significant split of the right branch (Node 9), with those belonging to MLU Group 2 producing more voice-marked verbs compared to those in MLU Group 1. Utterance type does not produce a significant split of the data.

The conditional variable importance scores from the conditional random forest model (Figure 4) show that participant ID is the most important predictor for voice inflection, similar to the first split in the conditional inference tree. This is followed by MLU group, and whether or not the utterance is imperative. Utterance type is deemed unimportant.

The models have satisfactory to good predictive power. The conditional inference tree has a classification accuracy of 0.76, while the C-index is 0.82. The

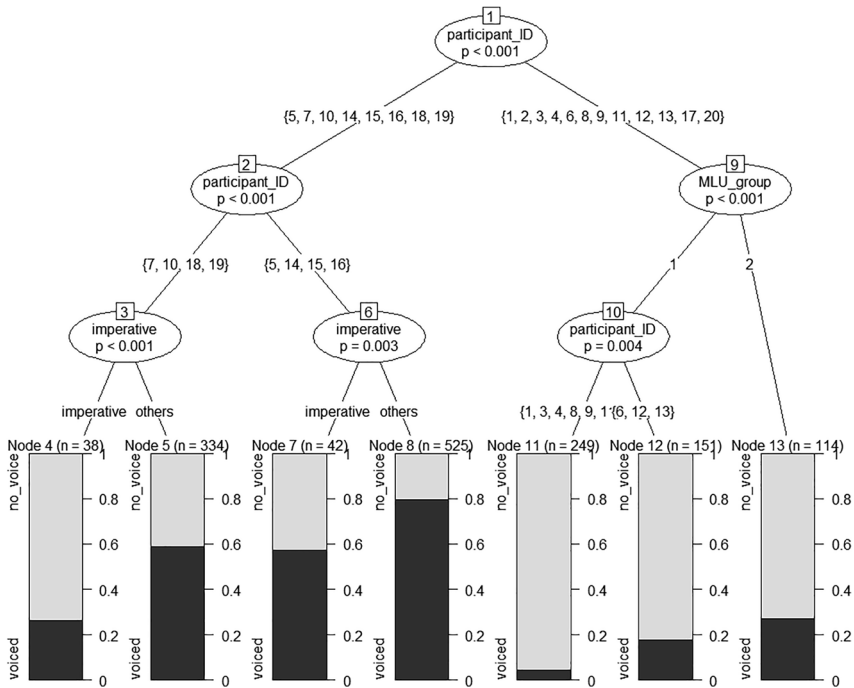


Figure 3: Conditional inference tree for voice inflection in children’s speech. Numbers on top of the circles indicate the node numbers. The *p*-values show the significance of the association between the independent variable in the circle and voice inflection. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, *n* indicates the number of observations, and the bar graphs indicate the fraction of voice-marked verbs.

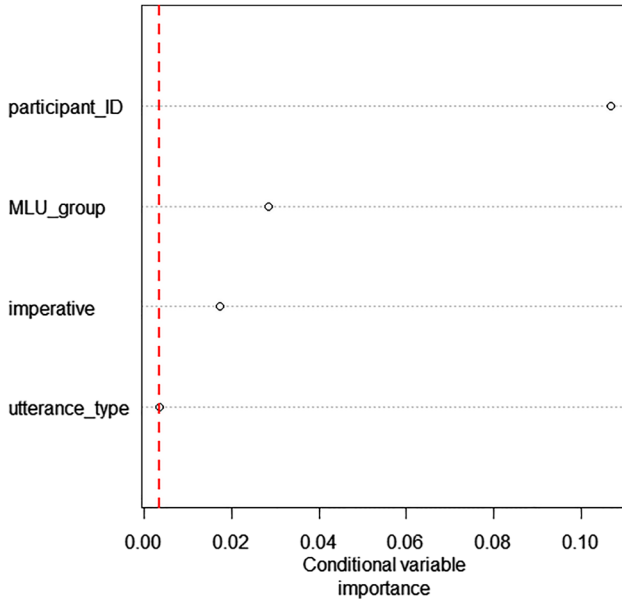


Figure 4: Conditional variable importance scores of the conditional random forest for voice inflection in children's speech. Predictors with dots to the right of the red vertical line are significant.

conditional random forest's out-of-bag classification accuracy is 0.77, and the out-of-bag *C*-index is 0.81.

The results show that whether or not children inflect the verb for voice is mostly due to individual variation, similar to the caregivers. Interestingly, the children who are more likely to mark voice on the verb (left branch of the first split of the tree) have caregivers who also mostly produce voice-marked verbs. Except for three children, those who are less likely to mark voice on the verb (right branch of the first split) have caregivers who inflect voice on the verb less. Although this is a purely descriptive result, this finding seems to show that the caregiver's use of voice-marking is related to children's verbal productions, as would be expected in approaches to acquisition that attribute high importance to the input (e.g., Lieven 2016; Tomasello 2003). There are two possibilities regarding this likely relationship: (i) the result we found from the caregivers is a stable property of the caregivers' child-directed speech, so their children are acquiring the unmarked forms, or (ii) the result could be particular to the recording session, such that children used the root words because the caregivers did.¹⁰ The latter means that children were primed to use the unmarked forms, and

¹⁰ Note that it is possible that the adults were primed by the children, which would be covered under possibility (i).

their productions are not representative of their competence. Our data do not allow us to decide between these two possibilities, but needless to say, the use of uninflected forms seems to be a prominent feature of the child-directed speech register and is replicated by children. Caregivers appear to simplify the input by using uninflected verbs, and this may have an influence on children's early productions. We also have evidence that language proficiency level has an influence on voice-marking, such that children who have higher MLU are more likely to inflect voice on the verb compared to those who have lower MLU. Additionally, similar to the caregivers, there is a tendency for imperatives to be voice-inflected less often than non-imperative utterances. Unlike in child-directed speech, we do not have evidence that verbs in simpler utterances (i.e., verb-initial) are inflected for voice less often than verbs in more complex utterances (i.e., non-verb-initial). Then again, children produced significantly fewer non-verb-initial sentences.

Overall, we found that for both caregivers and children, inflecting the verb for voice is mostly due to individual variation. We also found that imperative utterances tend to be inflected for voice less often than non-imperative utterances. Additionally, we have evidence that non-verb-initial utterances might carry more information for learning the voice-marking system compared to verb-initial utterances. Then again, they have non-canonical word order, and are typically more complex (involving clefts or relative clauses), and thus any added value they may have in helping children acquire the voice system may be offset by complexity. Lastly, we also found an influence of MLU group in children's productions, such that those with a higher MLU produce more inflected verbs compared to the group with a lower MLU.

3.1.2 Voice type

This analysis investigated the factors that determine the choice of agent and patient voice. We tested the following factors: utterance type (verb-initial or not), causativity (intransitive, transitive causative, transitive non-causative), presence of a definite non-agent argument (+/- definite non-agent), presence of an animate non-agent argument (+/- animate non-agent argument), and the individual variation among the speakers (participant ID). Utterances with non-voice-inflected verbs and utterances with verbs whose causativity could not be determined (from using filler words as verbs e.g., *nag-aano* 'what-ing') were excluded from the analyses. For children's speech, the influence of children's MLU group on choice of voice type was also tested.

Based on the literature on voice type alternation, we predicted that patient voice will be preferred over the agent voice in utterances depicting highly causative events (Katagiri 2005; Nolasco 2005), and in utterances with a definite non-agent argument (Himmelmann 2005b; Rackowski and Richards 2005) and/or an animate non-agent argument (Latrouite 2011). Given that different rules

supposedly obtain in relative clauses, existential and cleft constructions (Himmelmann 2005b), we predicted that the strong preference for the patient voice will be lessened in non-verb-initial constructions.

3.1.2.1 Child-directed speech

There were a total of 6,154 utterances with verb phrases which were marked for voice. Sixty-nine percent were in the patient voice and 31% were in the agent voice. The conditional inference tree on utterances with a voice-marked verb (Figure 5) shows that the first split comes from causativity. Intransitive utterances are more often in the agent voice than in the patient voice, while transitive causative and non-causative utterances are generally in the patient voice. This tendency of intransitive utterances to be in the agent voice is more strongly exhibited by non-verb-initial utterances without a definite non-agent argument compared to verb-initial utterances (Node 3). However, intransitive utterances with a definite

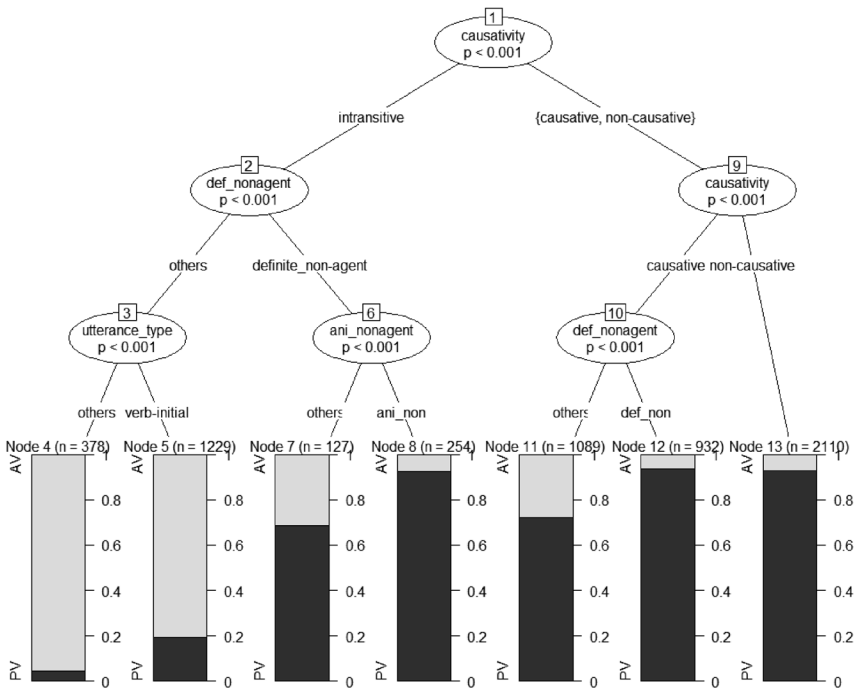


Figure 5: Conditional inference tree for voice type in child-directed speech. Numbers on top of the circles indicate the node numbers. The *p*-values show the significance of the association between the independent variable in the circle and voice type. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, *n* indicates the number of observations, and the bar graphs indicate the fraction of patient voice verbs.

non-agent show a strong tendency to be in the patient voice, which is more strongly exhibited by those with animate non-agents, compared to those without animate non-agents (Node 6).

Transitive utterances are further split, as the presence of a definite non-agent creates a split in causatives but not in non-causative utterances (Node 9). Under Node 10, transitive causative utterances with a definite non-agent argument show a stronger preference for the patient voice compared to those without a definite non-agent argument. Participant ID is not present in the tree, as it provides no significant splits of the data at the first three splits.

The conditional variable importance scores obtained from the conditional random forest model are in Figure 6. The results show that causativity is the most important predictor for voice choice, similar to the first split in the conditional inference tree. The second most important predictor is the presence of a definite non-agent argument. Other variables such as utterance type, the presence of an animate non-agent, and the participant IDs are not significant.

The models have good predictive power. The conditional inference tree has a classification accuracy of 0.86, while the concordance index C is 0.88. The conditional random forest's out-of-bag classification accuracy is 0.86, and the out-of-bag C is 0.90.

Most of our predictions for the factors influencing voice type alternation were borne out. The results show that the agent voice is the preferred voice for intransitive utterances, and the patient voice for both causative and

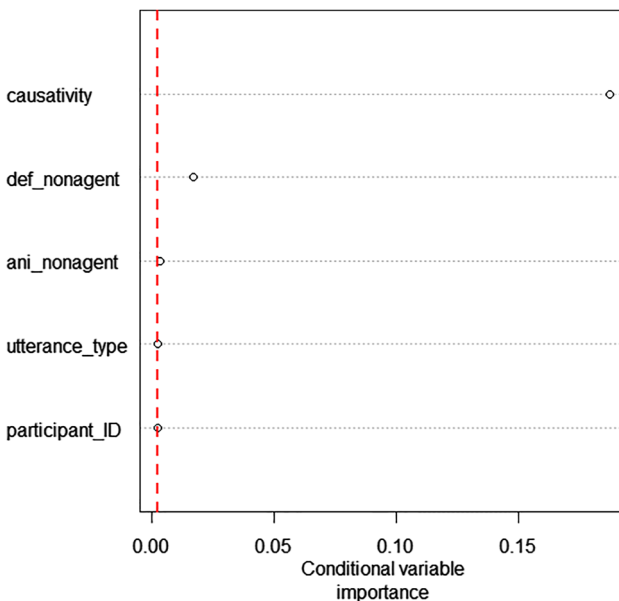


Figure 6: Conditional variable importance scores of the conditional random forest for voice type in child-directed speech. Predictors with dots to the right of the red vertical line are significant.

non-causative transitive sentences. This finding supports the claim that the patient voice is preferred for highly semantically transitive or causative events which involve a volitional agent and an affected non-agent argument (Katagiri 2005; Latrouite 2011; Nolasco 2005), similar to reports from written corpus analysis (Cooreman et al. 1984), and spontaneous child-directed speech analysis (Garcia et al. 2019). However, we also found that non-causative transitive verbs show a stronger preference for the patient voice compared to causative transitive verbs; which seems to indicate that transitivity (instead of causativity) is the most important predictor for voice type. Then again, many of the productions containing non-causative verbs which were mostly in the patient voice frequently occurred in almost the same constructions, e.g., *Anong ginagawa niya?* “What is he/she doing?” and *Tignan mo* “You look”. The inflected verb forms *ginagawa* and *tignan* actually make up 43% of all non-causative utterances. This effect may thus be most prominent in child-directed speech, since utterances such as these function to elicit conversation or direct attention, and are thus frequent components of caregiver-child interaction.

A deviation from the tendency for intransitive utterances to be in the agent voice seems to be only due to competition with the presence of a definite non-agent. Our findings support the claim that the patient voice is preferred given a definite non-agent argument (Himmelmänn 2005b; Latrouite 2011). We also have indications in the data that the patient voice is preferred whenever there is an animate non-agent argument, consistent with Latrouite’s and Himmelmänn’s claim. However, our evidence is only from the conditional inference tree and not from our conditional random forest analysis. Finally, there are indications (from the conditional inference tree but not from the conditional random forest) that there are differences between verb-initial and non-verb-initial utterances, with the latter showing less preference for the patient voice; which supports Himmelmänn’s claim that different rules for voice selection obtain in relative clauses, existential and cleft constructions, which we categorized under non-verb-initial utterances.

3.1.2.2 Children’s speech

Group 1 produced only 92 utterances with verb phrases which were marked for voice, while Group 2 had 625 utterances. Overall, 61% were in the patient voice and 39% were in the agent voice. The conditional inference tree on utterances with a voice-marked verb (Figure 7) shows that the first split is based on causativity. Intransitives do not show a preference for any voice, while transitive causatives and transitive non-causatives strongly favor the patient voice. For both branches (Nodes 2, 7), utterances with a definite non-agent argument show a higher tendency to be in the patient voice compared to those without a definite non-agent argument. Participant ID also has an influence on the left branch (Node 4), as a few

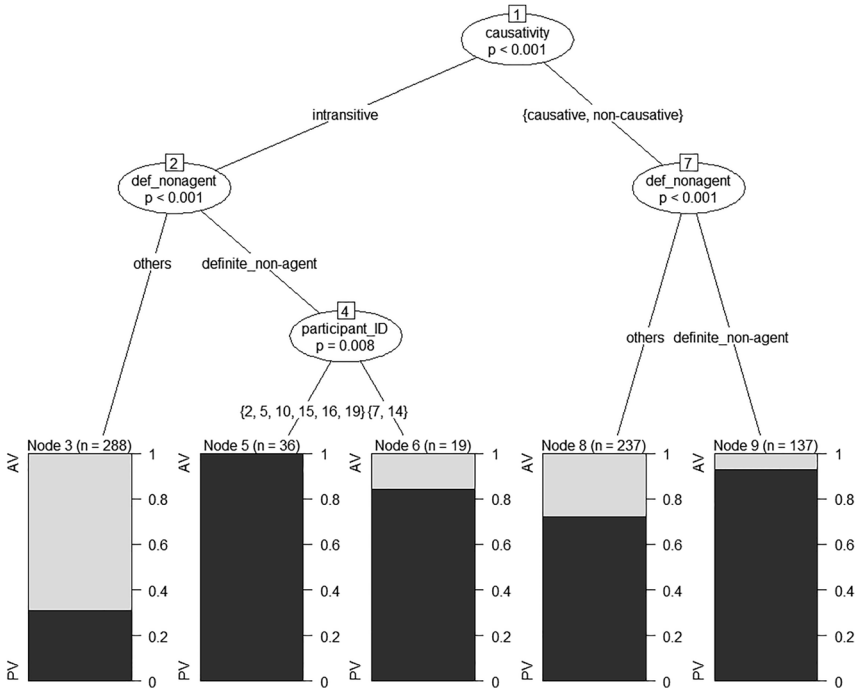


Figure 7: Conditional inference tree for voice type in children's speech. Numbers on top of the circles indicate the node numbers. The p -values show the significance of the association between the independent variable in the circle and voice type. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, n indicates the number of observations, and the bar graphs indicate the fraction of patient voice verbs.

participants produce intransitive utterances with definite non-agent exclusively in the patient voice. Other predictors (i.e., presence of an animate non-agent argument, utterance type, MLU group) are not present in the tree, as they provide no significant splits of the data.

The conditional variable importance scores obtained from the random forest model are in Figure 8. The results show that causativity is the most important predictor for voice choice, while the second most important predictor is the presence of a definite non-agent argument; similar to the results of the conditional inference tree. Participant ID is also of importance. The presence of an animate argument, utterance type, and MLU group are deemed unimportant.

The models have satisfactory predictive power. The conditional inference tree has a classification accuracy of 0.77, while the concordance index C is 0.80. The

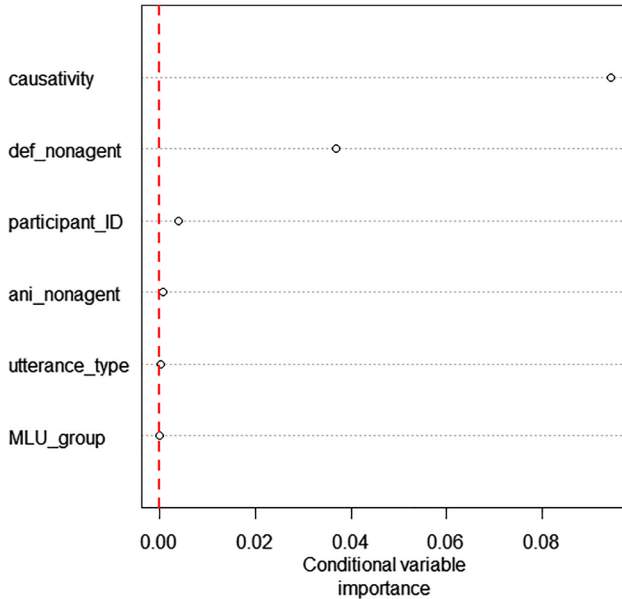


Figure 8: Conditional variable importance scores of the conditional random forest for voice type in children’s speech. Predictors with dots to the right of the red vertical line are significant.

conditional random forest’s out-of-bag classification accuracy is 0.77 and its out-of-bag C is 0.79.

The results show that, similar to our predictions based on the child-directed speech, children’s choice of voice-marking is dependent on causativity and the presence of a definite non-agent argument. Similar to caregivers, children prefer the patient voice given a definite non-agent argument, as well as for transitive causative and non-causative sentences. However, in children’s productions, intransitive utterances do not show a tendency for any voice, while for the caregivers, they tend to be in the agent voice. Thus, in this age range, we can say that children are still acquiring the nuances of how transitivity correlates with voice. Importantly, there is no significant influence of MLU group, although we have evidence for the influence of individual speaker variability.

These results suggest that children are able to use these language-specific factors as starting points for acquiring the complex system of voice-marking in Tagalog, namely that the patient voice should be used given a highly causative verb or if there is a definite non-agent argument. Even though children mostly use uninflected verbs at an earlier stage, the current results suggest that children’s emerging use of voice inflection is constrained in much the same way as in the input.

3.2 Word order and thematic role assignment

3.2.1 Word order

This analysis investigated the factors that influence the identity of the first NP, or in other words, the predictors which determine whether an argument will be realized as NP1 or not (here NP2 refers to all non-NP1 arguments, so it includes both arguments in NP2 and NP3 positions). We tested the following factors: utterance type (verb-initial or not), thematic role of the argument (agent, non-agent), nominality of the argument (monosyllabic pronoun, disyllabic pronoun, noun, question word), animacy of the argument (animate, inanimate, activity/referent), grammatical relation of the argument (*ang*-phrase or not an *ang*-phrase), and the individual variation among the speakers (participant ID). In the analysis of children's speech, MLU group was not included as a predictor because only data from Group 2 were analyzed, as explained in the Children's speech subsection below. Only transitive utterances with at least two arguments were included in the analyses. Utterances with the pronoun *kita* (I to you) were excluded from the analyses because two different thematic roles are coded as one morpheme, therefore occupying only one NP position. Clausal arguments (nominality of argument) were also excluded due to sparsity. Arguments whose animacy could not be identified were also excluded.

We predicted that agents (Kroeger 1993b; Riesberg et al. 2019), pronouns (Billings 2005; Reid and Liao 2004), and animate arguments (Tanaka et al. 2011) will most likely occupy the first argument position (NP1). We also predicted that *ang*-marked arguments will favor the NP2 position (Kroeger 1993b; Riesberg et al. 2019). We also investigated the role of utterance type (verb-initial or non-verb-initial) given that the linguistic theories on Tagalog argument order are specifically for verb-initial utterances (Riesberg et al. 2019).

3.2.1.1 Child-directed speech

Overall, there were 3,255 transitive utterances with no dropped argument. After excluding arguments with the pronoun *kita*, arguments whose animacy was labeled as 'unsure', as well as clausal and existential arguments (too few in comparison to others), there were 6,275 arguments. There were slightly fewer NP1 arguments (49%) than NP2 arguments (51%), as NP3s were considered as NP2.

The conditional inference tree on transitive utterances with complete arguments (Figure 9) shows that the first split comes from nominality, with monosyllabic pronouns and question words being mostly NP1, and disyllabic pronouns and nouns being mostly NP2. However, monosyllabic pronouns in non-verb-initial utterances favor the NP2 position (Node 8), and in verb-initial utterances, monosyllabic pronouns that are non-agents slightly favor the NP2 position (Node 5).

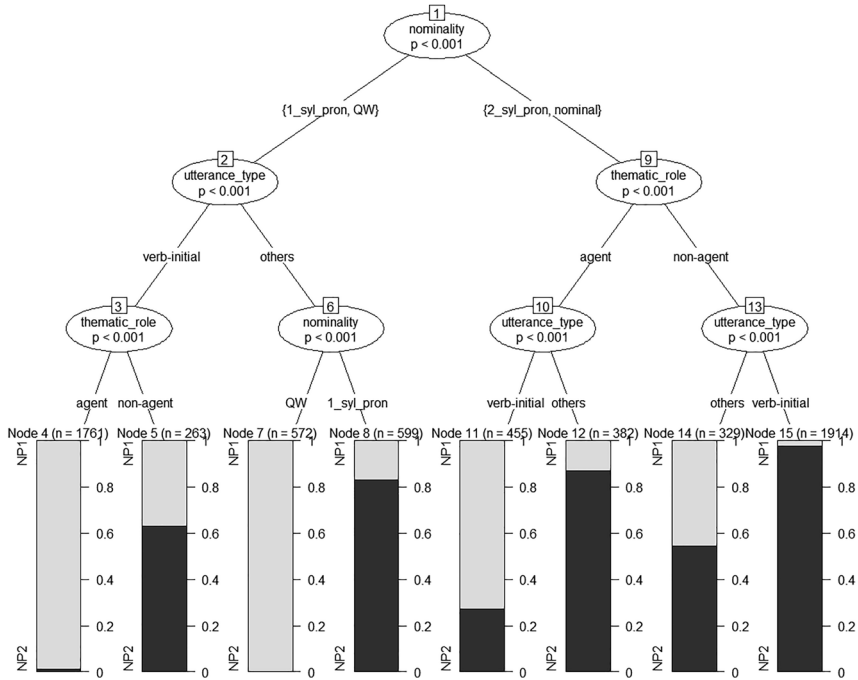


Figure 9: Conditional inference tree for word order in child-directed speech. Numbers on top of the circles indicate the node numbers. The *p*-values show the significance of the association between the independent variable in the circle and word order. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, *n* indicates the number of observations, and the bar graphs indicate the fraction of NP2s (second-mentioned arguments).

(Question words only appear in non-verb-initial utterances in the corpus). Additionally, the tree shows that nominal and disyllabic pronominal arguments occur more often as NP1 than NP2 when they are agents in verb-initial utterances (Node 11), and that these arguments have a lower tendency to be an NP2 when they are non-agents in non-verb-initial utterances compared to non-agents in verb-initial utterances (Node 13). Variables such as grammatical relation, animacy, and participant ID do not produce a significant split in the first three split levels.

The conditional variable importance scores from the conditional random forest model (Figure 10) show that utterance type is the most important predictor for word order, followed by thematic role and nominality. This is different from the result of the conditional inference tree, where the first split is based on nominality. However, utterance type appears three times in the first three levels of the tree, showing that it interacts with the other variables.

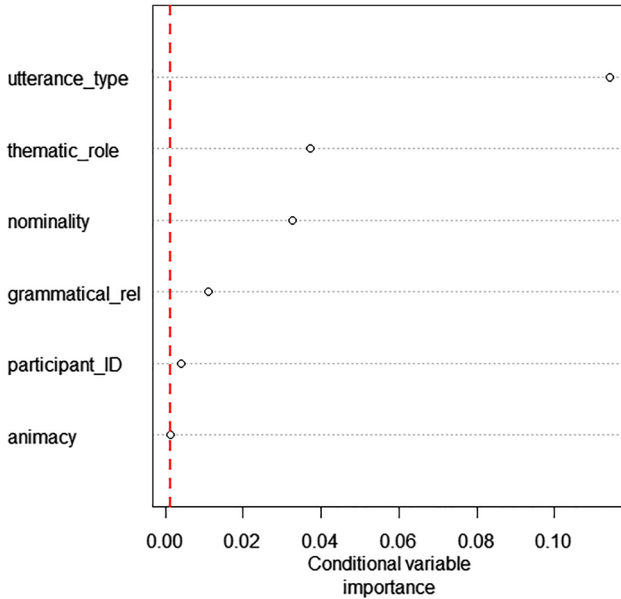


Figure 10: Conditional variable importance scores of the conditional random forest for word order in child-directed speech. Predictors with dots to the right of the red vertical line are significant.

Grammatical relation (*ang*-phrase or not) is also important in determining the argument position. Based on partial dependence plots,¹¹ *ang*-phrases tend to occur as NP1 more than non-*ang*-phrases, most likely because question words were considered as *ang*-phrases. However, the difference in the probability of the *ang*-argument and non-*ang*-argument to be the NP1 is small (0.54 for the *ang*-argument, 0.46 for the non-*ang*-argument; the average effects of the other predictors are accounted for). Participant ID is only marginally significant, and animacy does not seem to be important.

The models have excellent predictive power. The conditional inference tree has a classification accuracy of 0.91, while the concordance index C is 0.96. The conditional random forest's out-of-bag classification accuracy is 0.92, and the out-of-bag C is 0.98.

Interpreting the results of the conditional inference tree and conditional random forests together, we found that, in support of our predictions, monosyllabic pronouns and question words tend to be realized as NP1s, while disyllabic pronouns and nominal arguments tend to be NP2s. These results support claims on Tagalog that pronouns are mentioned as first arguments (Billings 2005; Schachter

¹¹ Partial dependence plots can be found in the study's OSF page: <https://osf.io/q568p/>.

and Otanés 1972), and that monosyllabic pronouns precede disyllabic pronouns (Reid and Liao 2004; Riesberg et al. 2019). The findings also show that agents tend to be realized as NP1s, while non-agents tend to be NP2s (see 13 and 14 for examples). This finding supports Kroeger's (1993b) proposal on the principles guiding the order of arguments in Tagalog, as well as Riesberg et al.'s (2019) argument for a universal agent-first preference.

However, we have evidence that nominality and thematic role also interact with utterance type, as monosyllabic pronouns in non-verb-initial utterances favor the NP2 position (see Example 15); and disyllabic pronominal and nominal agents in non-verb-initial utterances are mostly in the NP2 position (see Example 16). Both seem to be due to questions, and question words (in these cases, referring to a non-agent) usually come first. For the latter, two ordering principles compete with each other – agents tend to be NP1s while disyllabic pronominal and nominal arguments are usually NP2s. We found that the cues usually do not compete with each other, that is, 89% of monosyllabic pronouns were agents, and 64% of disyllabic pronouns and nominals were non-agents. Moreover, 73% of agents were produced as monosyllabic pronouns; and 49% of non-agents were produced as nominals and a further 24% of non-agents were produced as disyllabic pronouns.

In cases where the ordering principles do compete with each other, the principle that wins the competition seems to vary between utterance types. In non-verb-initial utterances, nominality of the argument seems to be more influential than thematic role, as disyllabic pronominal or nominal agents tend to occur as NP2. However, in verb-initial utterances, agency seems to be more influential than nominality, as agent nominal and disyllabic pronominal arguments occur mostly as NP1, and non-agent monosyllabic pronouns slightly favor the NP2 position. In general, these results suggest that word order (verb-initial versus non-verb-initial) is an important consideration when examining argument order in Tagalog.

- (13) *Lutu-in mo na ito.*
 cook-PV.INF 2.SG.NSBJ already 3.SG.SBJ
 'You cook this already.' (C14)
- (14) *Lutu-in mo 'yung isda.*
 cook-PV.INF 2.SG.NSBJ SBJ fish
 'You cook the fish.' (C20)
- (15) *Ano-ng ni-lu~luto mo?*
 what-LIN PV-IPFV~cook 2.SG.NSBJ
 'What are you cooking?' (C2)
- (16) *Ano lu~lutu-in natin?*
 what CONT~cook-PV 1.PL.NSBJ
 'What are we going to cook?' (C3)

Our findings on grammatical role are contrary to our predictions. We found that *ang*-arguments are more likely to occur as NP1, whereas we predicted that they would tend to occur as NP2; although the difference with non-*ang*-arguments is rather small. Inspecting the data shows that this seems to be mostly due to the high number of questions. Question words usually occur as NP1, and since we marked question words as *ang*-arguments, there are many examples of the *ang*-argument occurring as the first argument. Other instances of the *ang*-phrase occurring as NP1 include those where the *ang*-phrase is a monosyllabic pronoun and the non-*ang*-phrase is a noun; showing that the nominality constraint on word order is strong. Others include fronting the *ang*-argument. It could be the case that the *ang*-last order applies only to verb-initial utterances with full noun phrases, and we do not observe this in the analysis, as we included non-verb-initial utterances, and marked question words and *ang*-form pronouns as *ang*-arguments.

We also predicted that animate arguments will be mentioned first, following crosslinguistic claims that animates tend to precede inanimates (Ferreira 1994; Minkoff 2000; Van Valin and LaPolla 1997). However, we did not find evidence for an influence of animacy on word order. Perhaps this is because we also included utterances with two animate arguments in our counts (38% of utterances which qualified for the word order analysis), resulting in only 56% of animates occurring as NP1 (and the rest of the 44% as NP2). However, 99% of agent arguments were actually animate, and 97% of monosyllabic pronominal arguments were also animate; so the influence of animacy is likely covered by the thematic role and nominality predictors.

Overall, word order or argument position seems to be mostly determined by utterance type, thematic role, and nominality, which implies that these cues are available to children for learning Tagalog's word order patterns. Additionally, since animacy usually overlaps with thematic role and nominality (i.e., agents and pronouns are usually animate), animacy might also be a useful cue for children to learn the word order patterns of Tagalog.

3.2.1.2 Children's speech

Group 1 produced only 19 transitive utterances with no dropped argument. In total, there were 39 arguments. The majority of the arguments were pronouns (54% monosyllabic pronouns, 28% disyllabic pronouns), and the remaining 18% were nouns. Sixty-four percent of the arguments were animate, and the other 36% were inanimates. Given this low number, we only analysed Group 2's productions. Group 2 produced 188 transitive utterances with no dropped argument. After excluding those with the pronoun *kita*, arguments which were labeled 'unsure' for animacy, and the 17 arguments which were marked as question word or clause, 368

arguments were left. There were slightly fewer NP1 arguments (48%) than NP2 arguments (52%) as NP3s were considered as NP2.

The conditional inference tree on Group 2’s transitive utterances with complete arguments (Figure 11) shows that the first split comes from thematic role, with agents being generally realized as NP1 and non-agents being frequently realized as NP2. Furthermore, Node 9 shows that inanimate non-agents in verb-initial utterances show a stronger tendency to occupy the NP2 position compared to animate non-agents. However, non-agents in non-verb-initial utterances show no tendency for any argument position (Node 8). Moreover, the tree also shows that agents in non-verb-initial utterances slightly favor the NP2 position (Node 3), and nominal agents in verb-initial utterances also favor the NP2 position (Node 5). Variables such as grammatical relation (*ang*-phrase or not) and participant ID do not produce a significant split of the data.

The conditional variable importance scores from the conditional random forest model (Figure 12) show that similar to the conditional inference tree, thematic role is

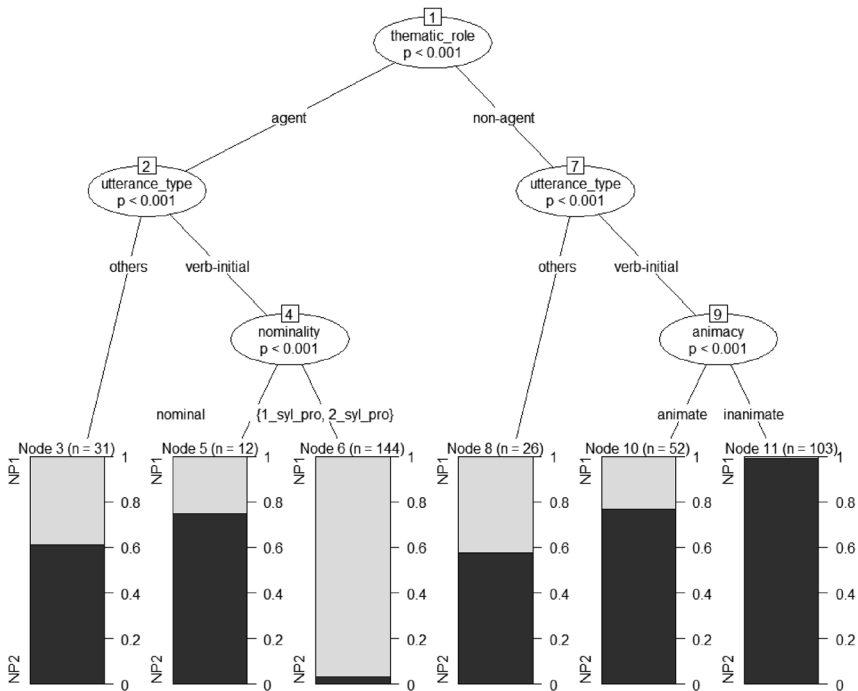


Figure 11: Conditional inference tree for word order in children’s speech. Numbers on top of the circles indicate the node numbers. The *p*-values show the significance of the association between the independent variable in the circle and word order. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, *n* indicates the number of observations, and the bar graphs indicate the fraction of NP2s (second-mentioned arguments).

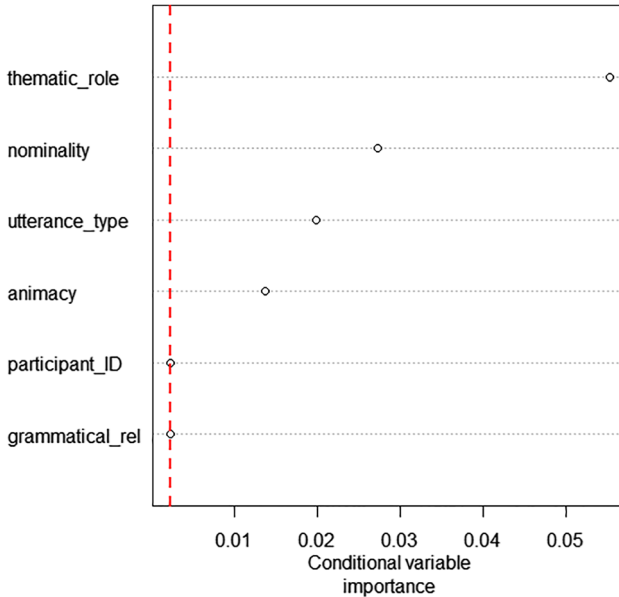


Figure 12: Conditional variable importance scores of the conditional random forest for word order in children’s speech. Predictors with dots to the right of the red vertical line are significant.

the most important factor. This is followed by nominality, utterance type, and animacy. Grammatical relation and participant ID are deemed unimportant.

The models have good to excellent predictive power. The conditional inference tree has a classification accuracy of 0.88, while the concordance index C is 0.94. The conditional random forest’s out-of-bag classification accuracy is 0.87, and the out-of-bag C is 0.94.

Interpreting the results of the conditional interference tree and conditional random forests together shows that children, at least those with a higher MLU, respect both crosslinguistic and language-specific constraints on argument position. Similar to caregivers, Group 2 children’s word order or argument position seems to be mostly determined by thematic role, nominality, and utterance type. In general, agents favor the NP1 position, while non-agents favor the NP2 position; similar to previous experimental (Bautista 1983; Tanaka 2016) and corpus data findings (Garcia 2016). A deviation from this pattern seems to be due to the competition between thematic role and nominality, as the findings also show that pronominal agents in verb-initial utterances tend to be realized as NP1s, while nominal agents tend to be NP2s. These findings are in line with argument ordering proposals for Tagalog (Kroeger 1993b; Riesberg et al. 2019 for agent-first order;

Billings 2005; Schachter and Otnes 1972 for a pronoun-first order), as well as with supposedly universal linearization hierarchies on thematic roles and referentiality (see Siewierska 1993 for a review). Utterance type also seems to interact with thematic role: agents in non-verb-initial utterances slightly favor the NP2 position, and non-agents in non-verb-initial utterances show no tendency for any position. It must be mentioned though that arguments which occurred in non-verb-initial utterances comprised only 15% of Group 2's productions.

However, in contrast to child-directed speech, animacy also plays a role in children's word order choice. Children tend to produce the animate argument as the NP1 and the inanimate argument as NP2, which is in line with the claim that animates tend to precede inanimates universally (Ferreira 1994; Minkoff 2000; Van Valin and LaPolla 1997). Similar with the child-directed speech, animacy usually does not go against thematic role nor nominality, that is, agents and monosyllabic pronouns are mostly animates, and non-agents and nominals are more often inanimates than animates.

3.2.2 Thematic role of the NP1

This analysis investigated the factors that influence the thematic role of the first argument (i.e., agent, non-agent). We limited our investigation on the NP1 so we could also include intransitives in our analysis, for which there is no prior data on Tagalog. We tested the following factors: utterance type (verb-initial or not), nominality of the first argument (monosyllabic pronoun, disyllabic pronoun, noun, question word), animacy of the first argument (animate, inanimate, activity/referent), voice-marking on the verb (agent voice, patient voice), and the individual variation among the speakers (participant ID). In the analysis of children's speech, MLU group was not included as a predictor because only data from Group 2 were analyzed, as explained in the Children's speech subsection below. Only utterances with at least one argument were included in the analyses. Similar to argument position models, utterances with the pronoun *kita* (I to you), clausal arguments (nominality of argument), and arguments whose animacy could not be identified were excluded from the analyses. Additionally, utterances with verbs which were not voice-marked were excluded. We expected that NP1s that are animate (Dowty 1991; Hopper and Thompson 1980) and pronominal (Croft 2003; Silverstein 1976) are most likely the agents. Additionally, we predicted that NP1s in the agent voice will more likely occur as agents compared to NP1s in the patient voice (Garcia et al. 2019).

3.2.1.1 Child-directed speech

In the 5,215 utterances qualified for the analysis, 71% of NP1s were agents. The conditional inference tree on utterances with at least one argument (Figure 13)

shows that the first split comes from animacy, with animate NP1s being mostly agents, and inanimate and activity/referent NP1s occurring mostly as non-agents. Pronominal animate NP1s also have a stronger tendency to be agents in the agent voice than in the patient voice (Node 10). Furthermore, question words referring to activity or inanimate arguments occur exclusively as non-agents (Node 8). However, Node 4 shows that activity or inanimate arguments coded by pronouns in the agent voice do not show a tendency for any thematic role. Additionally, animate NP1s coded by nouns or question words in the patient voice show a slight preference for non-agent roles (Node 15). Utterance type and participant ID do not produce a significant split in the first three split levels.

The conditional variable importance scores from the conditional random forest model (Figure 14) show that animacy is the most important predictor for the NP1's thematic role, similar to the first split in the conditional inference tree. The second most important predictor is nominality, followed by voice. Utterance type and participant ID are deemed unimportant.

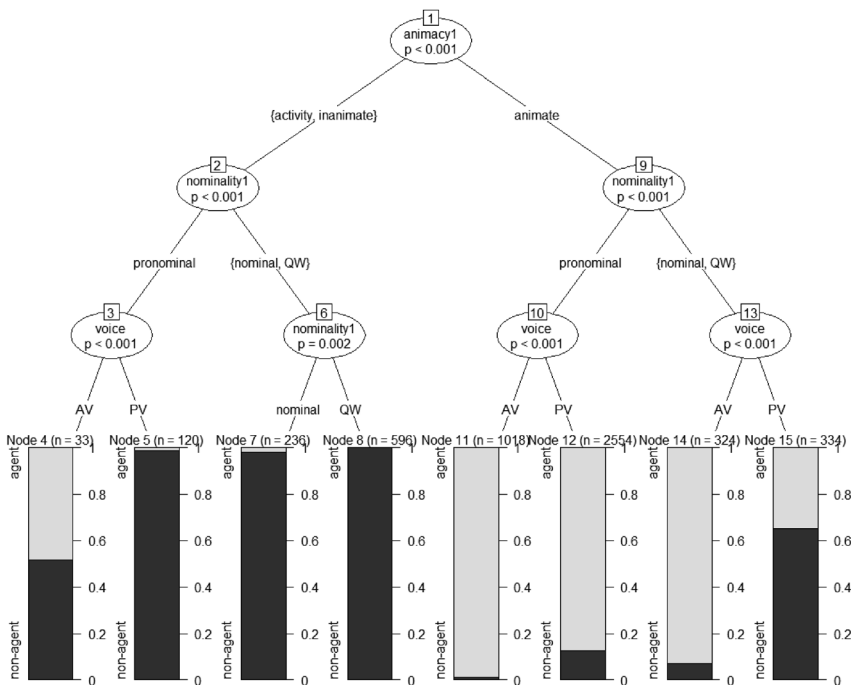


Figure 13: Conditional inference tree for thematic role of NP1 in child-directed speech. Numbers on top of the circles indicate the node numbers. The p -values show the significance of the association between the independent variable in the circle and thematic role. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, n indicates the number of observations, and the bar graphs indicate the fraction of non-agent arguments.

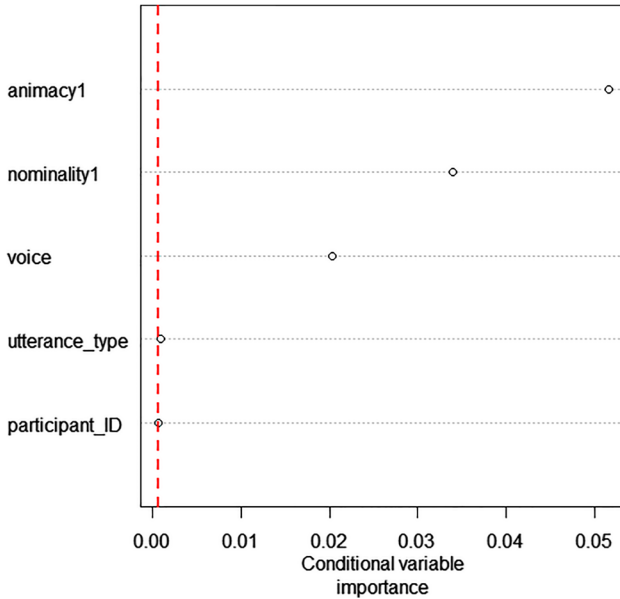


Figure 14: Conditional variable importance scores of the conditional random forest for thematic role of NP1 in child-directed speech. Predictors with dots to the right of the red vertical line are significant.

The models have excellent predictive power. The conditional inference tree has a classification accuracy and concordance index C of 0.91. The conditional random forest's out-of-bag classification accuracy is 0.90, and the out-of-bag C is 0.92.

The results show that similar to our predictions, animate NP1s tend to be agents, while inanimate and activity/referent NP1s tend to be non-agents. Moreover, pronominal NP1s have a stronger tendency to be agents compared to question words and nominal arguments, similar to what was reported by Garcia et al. (2018). These findings support universal claims on the influence of animacy and nominality on thematic roles (Croft 2003; Dowty 1991; Silverstein 1976). Additionally, we found that NP1s in the agent voice are more likely to occur as agents compared to NP1s in the patient voice; which is in line with findings on child-directed speech data which included only verb-initial utterances with highly causative verbs (Garcia et al. 2019). In the current data set, this finding seems to be due to the high frequency of *ano* 'what' questions in the patient voice, and since the question word refers to the non-agent in the patient voice, there were many instances of non-agent-initial utterances in this voice.

These results show that animacy and nominality of the first noun can serve as reliable cues for children to learn thematic role assignment. Moreover, regardless

of transitivity or whether or not the utterance is verb-initial, it seems that the first argument of agent voice utterances could be reliably interpreted as the agent.

3.2.1.2 Children’s speech

In Group 1’s utterances with at least one argument (180), 67% of NP1’s were agents. Seventy-eight percent of NP1s were animate, and also 78% of NP1s were pronominal. Because we wanted to test the influence of voice type on the thematic role of the first argument, we did not include Group 1’s productions in the models, as only 25% (46) of their verbs were inflected for voice.

In Group 2, of the 406 utterances that qualified for the analysis, 65% were in the patient voice. The majority of the NP1s were agents (71%), animates (86%), and pronominal (81%). The conditional inference tree on utterances with at least one argument (Figure 15) shows that the first split comes from animacy, with animate NP1s occurring frequently as agents, and inanimate NP1s occurring mostly as non-

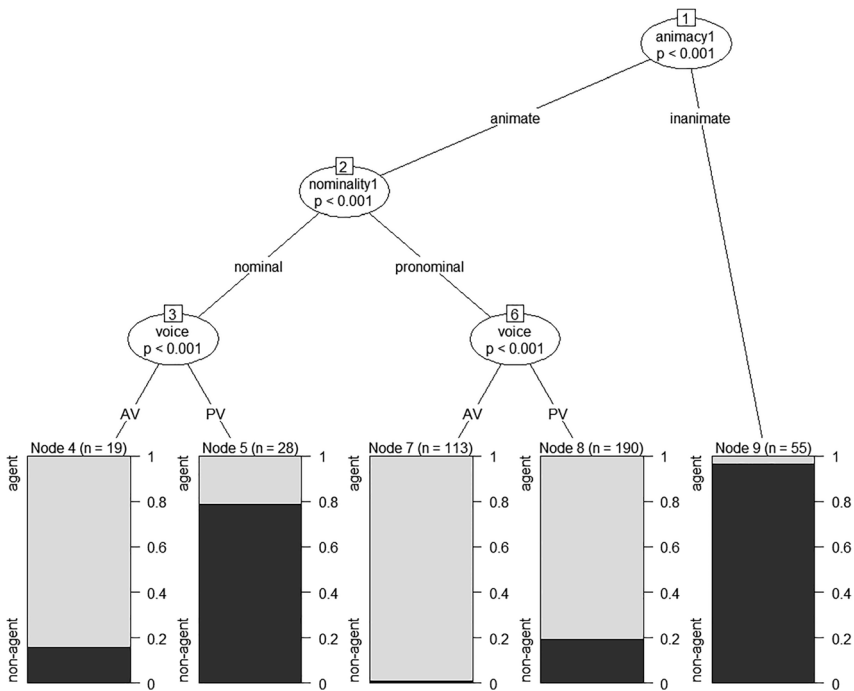


Figure 15: Conditional inference tree for thematic role of NP1 in children’s speech. Numbers on top of the circles indicate the node numbers. The *p*-values show the significance of the association between the independent variable in the circle and thematic role. The subsets created by the binary splits can be found below the circles, and are enclosed in brackets in case of multiple levels. In each terminal node, *n* indicates the number of observations, and the bar graphs indicate the fraction of non-agent arguments.

agents. Pronominal animate arguments also show a higher tendency to be agents compared to nominal animate arguments (Node 2). Additionally, pronominal animate arguments in the agent voice almost exclusively occur as agents (Node 7). However, nominal animate arguments in the patient voice favor the non-agent role (Node 5). Utterance type and participant ID do not produce a significant split of the data.

The conditional variable importance scores from the conditional random forest model (Figure 16) show that animacy is the most important predictor for the NP1's thematic role, similar to the first split in the conditional inference tree. The second most important predictor is nominality, followed by voice. Utterance type and participant ID are deemed unimportant.

The models have good predictive power. The conditional inference tree has a classification accuracy and concordance index C of 0.88. The conditional random forest's out-of-bag classification accuracy and out-of-bag C are both 0.88.

The results are very similar to those of the caregivers. Animate NP1s tend to be agents, while inanimate NP1s tend to be non-agents. Additionally, pronominal animate NP1s have a stronger tendency to be agents compared to nominal NP1s.

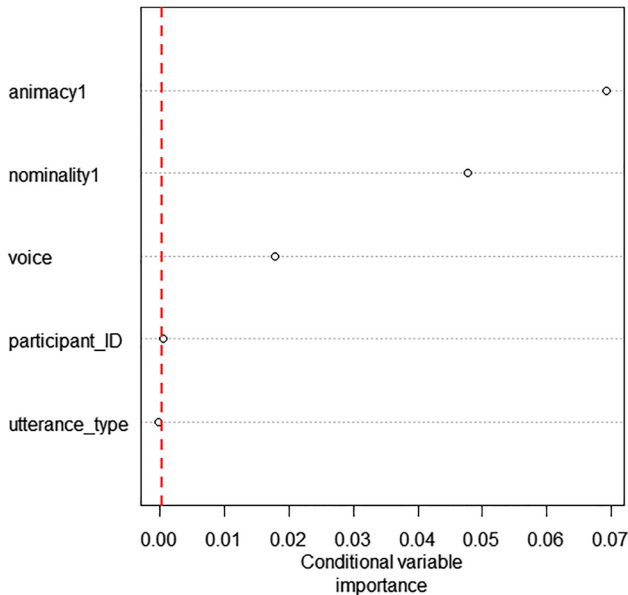


Figure 16: Conditional variable importance scores of the conditional random forest for thematic role of NP1 in children's speech. Predictors with dots to the right of the red vertical line are significant.

These findings reveal that children follow supposedly universal tendencies on the influence of animacy and nominality on thematic roles (Croft 2003; Dowty 1991; Silverstein 1976), similar to their caregivers. Moreover, animate NPs in the agent voice have a stronger tendency to be agents compared to NPs in the patient voice. In other words, agent voice sentences are more often agent-initial or agent-only, compared to patient voice sentences. This is partly in line with Garcia's (2016) finding that in the agent voice, children produced only 3% *non-agent-only* or *non-agent-initial* constructions, but in the patient voice, it was 6%.

4 Conclusions

In the current paper, we reported multivariate analyses of Tagalog child-directed and children's speech, in which we tested how multiple variables condition voice-marking and word order selection. To our knowledge, this is the first study of its kind in Philippine-type symmetrical voice language (see McDonnell 2016 and Riesberg et al. 2021, for voice selection in non-Philippine-type symmetrical voice languages based on analyses of adult speech), and the first to quantitatively assess these claims. As such, our study contributes both to the descriptive linguistics literature on Tagalog (following arguments that child language data is a legitimate source of documentary data, Hellwig and Jung 2020), and to the child language literature in novel ways. Regarding the linguistics of Tagalog voice and argument structure, we summarize our main findings in Table 2.

Our study adds to a small literature that has investigated variable voice selection in Austronesian languages. The finding that the patient voice is preferred given highly causative or semantically transitive events is in line with the findings in Totoli (Riesberg et al. 2021) and Besemah (McDonnell 2016). Additionally, the finding that the presence of an animate non-agent is linked to a preference for the patient voice in Tagalog is similar to the findings on Besemah. However, this is not fully in line with the findings on Totoli, which show that the patient voice is preferred regardless of the animacy of the non-agent argument. Then again, our statistical analyses did not test the influence of the presence of an inanimate non-agent nor of the interaction of the animacy of the agent and that of the non-agent on voice selection (e.g., voice tendency when only either the agent or the non-agent is animate), but merely the presence of an animate non-agent in an utterance. We also did not test several of the predictors tested in Totoli and Besemah (e.g., topicality, grounding, collostruction strength). This small sample suggests possible language-specific variation in the variables influencing voice-marking within the Austronesian language family, and more studies investigating variation in this domain would be welcome.

Table 2: Theoretical claims on voice-marking, word order selection, and thematic role assignment which are supported by the current findings. Those in bold are strongly supported while those which are not in bold are partially supported.

Predictors	References	Claims
Voice alternation		
Causativity	Katagiri 2005; Latrouite 2011; Nolasco 2005	Patient voice is preferred for highly causative (semantically transitive) events.
Definite non-agent	Himmelmann 2005b; Latrouite 2011; specificity of the non-agent argument: Adams and Manaster-Ramer 1988; Maclachlan and Nakamura 1997; Rackowski and Richards 2005	Patient voice is preferred in constructions with definite non-agent arguments.
Animate non-agent argument	Himmelmann 2005b; Latrouite 2011	Patient voice is preferred in constructions with animate non-agent arguments.
Utterance type	Himmelmann (2005b)	Different rules obtain in relative clauses, existential and cleft constructions compared to verb-initial utterances.
Word order		
Thematic role	Kroeger 1993b; Riesberg et al. 2019	The agent tends to be the first argument.
Nominality	Billings 2005; Reid and Liao 2004; Riesberg et al. 2019; Schachter and Otnes 1972	A pronoun is mentioned before a noun, and a monosyllabic pronoun is mentioned before a disyllabic pronoun.
Thematic role assignment		
Animacy	Dowty 1991; Hopper and Thompson 1980	An animate argument is more likely to be an agent than an inanimate argument.
Nominality	Croft 2003; Silverstein 1976	A pronominal argument is more likely to be an agent than a nominal argument.

With respect to child language, our findings show that despite the diversity of structures that children have to learn under Tagalog's voice system, there are factors that strongly predict the speakers' choice between the voice and word order alternations, which children can use as starting points for acquiring the voice system. The results also show that these tendencies usually do not compete against each other, so the cues available in the input seem to be reliable overall. For example, the deviations from the observation that intransitive verbs are in the

agent voice are mostly due to the presence of definite non-agents. However, intransitives more commonly do not have definite non-agents. For word order and thematic role assignment, thematic role, animacy, and nominality usually overlap, such that most agents are also typically animate and monosyllabic pronouns. More importantly, we found that children's productions seem to be guided, with some exceptions (mainly concerning choice of voice inflection and voice type), by the same factors that caregivers use for choosing between variations. This suggests that, early on, children's choices related to structure alternations are similar to what is available in their input, and are consistent with the caregiver model.

A final comment is warranted about the role of this study in the development and evaluation of language acquisition theory. The kinds of analyses that we have presented are very different from typical child language acquisition studies, due both to the nature of our corpus and our analytic strategy. Intensively sampled longitudinal corpora of monolingual Tagalog-speaking children do not exist, and so our approach was to collect semi-naturalistic samples using a cross-sectional method. This limits our ability to test theoretical models of acquisition, which typically asks questions like "when do children have abstract knowledge of category X?", most often on the basis of longitudinal data (e.g., Lieven et al. 2003). We did not attempt to ask these questions because of the nature of our corpus. However, we believe that our analytic approach has complementary value to traditional approaches, with the main strength being that our variationist approach enabled us to: (i) identify key variables that condition variation in the input, and (ii) test whether children are sensitive to these variables in their own speech. By and large, we found that children are sensitive to many of the same variables that condition variation in their input. This by no means suggests that they have full command of the language; however, what it does mean is that the pattern of data is compatible with the idea that they are using those cues to guide their Tagalog use (and presumably, their parsing of the input). Studies with denser longitudinal data would be able to test how these sensitivities change across development.

Thus, we see great utility in applying the variationist approach to child language acquisition. A common feature of all language acquisition theories is a commitment to distributional analysis in learning language-specific components from the input (e.g., Lidz and Gagliardi 2015; Saffran and Kirkham 2018; Tomasello 2003). However, studies of the acquisition of grammar, including verb argument structure, have rarely comprehensively analysed variables that condition variation, which might provide children with clear cues to break into the system (one notable exception on the English passive is Budwig 1990). Children acquire the linguistic practices of their community, and if we are to understand this process at a deeper level, we need to identify those variables that influence the distributions

that categorise those practices and to test if children show sensitivity to these variables. This requires moving beyond variables like frequency (Ambridge et al. 2015), to instead study how constraints from the language interact with their availability to jointly determine acquisition trajectories.

Acknowledgments: We would like to thank the families in Bagong Barrio, Caloocan City who welcomed us into their homes. We are also grateful to Angela Legaspi and Micah Marcial for transcribing and annotating the videos, and to Caroline Rowland and two anonymous reviewers for giving us feedback on an earlier draft. We also thank several people for sharing their expertise with us: Natalia Levshina for the statistical analysis, Marisa Casillas and Han Sloetjes for ELAN; Jeroen Geerts for video creation; and Ivan Bondoc, Jed Pizarro-Guevara, Ronel Laranjo, and Elsie Or for Tagalog.

Research funding: This work was supported by the Max Planck Society and the Australian Research Council (CE140100041: CI Kidd).

Appendices

Appendix A: Non-verb-initial Tagalog sentences

Inversion or fronting constructions include the placement of an *ang*-phrase before a verb, adverb before a verb or a verb complement before the verb. Inversion can involve *ay*-constructions (17a), the use of a pause after the fronted phrase (17b), initial placement of an adverb or direction that is not separated from the verb by a pause (17c), and initial placement of an adverb and insertion of a linker *ng/na*- before the first word of the predicate (17d).

- (17) a. *Ang bata ay t<um>alon*
 SBJ child <AV>jump.PFV
 ‘The child jumped.’
- b. *Si Papa, nag-re~relax.*
 SBJ Papa AV-IPFV~relax
 ‘Papa is relaxing.’ (C1)
- c. *Dito mo i-lagay.*
 here 2.SG.NSBJ PV-put.INF
 ‘You put (it) here.’ (C15)
- d. *Bigla siya-ng t<um>alon.*
 suddenly 3.SG.SBJ-LIN <AV>jump.PFV
 ‘He/She suddenly jumped.’ (C17)

Another form of fronted constructions are *wh*-questions (Sabbagh 2011). It has been further argued that question words that ask for an argument ([18a] compared to [18b] which does not ask for an argument) are pseudo-clefts (Aldridge 2002), where the *wh*-phrase is the predicate and the rest of the clause (a headless relative clause) is the subject. Formation of such *wh*-questions involves voice, as the question words can only refer to *ang*-arguments (Pizarro-Guevara and Wagers 2020). This entails that *sino* ‘who’ in an agent voice construction asks for the agent, while *sino* in a patient voice construction asks for a non-agent. In this paper, we grouped all *wh*-questions that ask for an argument as non-verb-initial sentences.

- (18) a. *Ano ang bi~bilh-in ni CHI*
 what SBJ CONT~buy-PV NSBJ child’s name
 ‘What is CHI going to buy?’ (C3)
- b. *Bakit siya la~labas?*
 why 3.SG.SBJ <AV>CONT~go_out
 ‘Why is he/she going out?’ (C5)

Other non-verb-initial structures involve negations (19). Tagalog also has the negative word *huwag* ‘don’t’ (20a) but we counted these under modals (e.g., *gusto* ‘want’ (20b), *ayaw* ‘don’t want’ (20c), *dapat* ‘should’, *pwede* ‘may’), which Schachter and Otones (1972) have labeled as pseudoverbs. It must also be mentioned that *gusto* and *ayaw* could have an agent voice-marked main verb with no other *ang*-phrase, so the main verb takes the *ng*-phrase following the *gusto/ayaw* phrase as its agent (20b), (20c). Reid and Liao (2004) refer to these as constructions with non-auxiliary extension verbs. This means that the agent in agent voice constructions with *gusto* and *ayaw* could be *ng*-marked.

- (19) *Hindi niya ma-huli ‘yung manok*
 NEG 3SG.NSBJ PV-catch.INF SBJ chicken
 ‘He/She cannot catch the chicken.’ (C20)
- (20) a. *Huwag mo-ng i-hagis*
 NEG 2SG.NSBJ-LIN PV-throw.INF
 ‘Don’t throw (it).’ (C17)
- b. *Gusto mo mag-basketball?*
 want 2SG.NSBJ AV-play_basketball.INF
 ‘Do you want to play basketball?’ (C2)
- c. *Ayaw mo pa-kingg-an ‘to*
 NEG 2SG.NSBJ CAU-listen-PV SBJ
 ‘You don’t want to listen to this.’ (C20)

We categorized the Tagalog negative word *wala* ‘none’ as part of existentials along with *may* and *mayroon* ‘there is’ (21a) and (21b). In verbal indefinite constructions, these words are equivalent to English indefinite pronouns (Schachter and Otanes 1972). The verbal phrases following existential and non-existential words have also been claimed to be relative clause modifiers (Sabbagh 2009).

- (21) a. *Wala ka naman-g i-lu~luto dito*
 NEG.EXIST 2SG.SBJ PV-CONT~cook here
 ‘You have nothing to cook here.’ (C16)
- b. *May nag-si~swimming*
 EXIST AV-IPFV~swimming
 ‘Someone is swimming.’ (C13)

Lastly, non-verb-initial Tagalog constructions include relative clause constructions (22) and cleft constructions (23). Similar to *wh*-question formation, voice interacts with relative clauses as only the *ang*-phrase can be relativized (Pizarro-Guevara and Wagers 2020). Relative clauses can be head-initial, head-final, internally headed or headless (Law 2016). A headless relative clause is one of the two *ang*-phrases that are juxtaposed in a cleft sentence (Kroeger 1993b).

- (22) *Halu-halu-in mo daw ‘yung ni-lu~luto mo*
 mix-PV.INF 2SG.NSBJ SBJ PV-IPFV~cook 2SG.NSBJ
 ‘Mix what you are cooking.’ (C5)
- (23) *Siya nga ‘yung g<um>a~gawa ng bahay*
 3.SG.SUBJ SBJ <AV>IPFV~build NSBJ house
 ‘She/He is the one building a house.’ (C7)

Appendix B: Voice affixes

Table 3 shows the voice, aspect, mood paradigm for the verb *kain* ‘eat’ based on Himmelmann (2005b). Regarding aspect and mood, non-realis/perfective refers to the infinitive, which is also the imperative in Tagalog. Non-realis/imperfective refers to an event that has not yet occurred (contemplated). Realis perfective refers to a finished event, while realis imperfective describes events that are ongoing, repeated, or habitual regardless of the time reference. The imperfective aspect is marked by consonant-vowel (CV) reduplication, while the perfective is unmarked. In agent voice *-um-* verbs, mood is distinguished only in the imperfective, as *-um-* is absent in non-realis mood. In agent voice *mag-* or *maN-* verbs, realis mood

Table 3: Dynamic voice, aspect, mood paradigm for *kain* ‘eat’ based on Himmelmann (2005b).

	Agent voice	Patient voice <in>	Patient voice –an	Patient voice i-/ipag-
Non-realis/Perfective	k<um>ain	kain-in	kain-an	i-kain
Non-realis/Imperfective	ka~kain	ka~kain-in	ka~kain-an	i-ka~kain
Realis Perfective	k<um>ain	k<in>ain	k<in>ain-an	i-k<in>ain
Realis Imperfective	k<um>a~kain	k<in>a~kain	k<in>a~kain-an	i-k<in>a~kain

Recent perfective: *ka-ka~kain*

is expressed by changing the voice markers *mag-* and *maN-* to *nag-* and *naN-*, respectively. In the patient voice, realis mood is marked by the infix *-in-*, and the suffix *-in* is dropped.

Tagalog also has a recent perfective aspect which refers to actions which have just occurred. The recent perfective has no voice-marking. It is marked only by the prefix *ka-* and the reduplication of the first syllable of a verb, e.g., *kakakain* ‘just ate.’ Moreover, the agent in recent perfective constructions cannot be marked by *ang*.

Additionally, dynamic verbs with a volitional or causative agent also have a potentive counterpart (see Table 4). The potentive is used for actions that are involuntary (‘I accidentally broke the glass.’) or with an unintended outcome (‘I sent the wrong documents’). It can also be used to express an achievement of the agent (‘I passed the exam.’). In our analysis, we did not distinguish the potentive use of *ma-/na-* and *maka-/naka-* from their stative use (to express the state of an entity). Instead, we counted all the potentive and stative *maka-/naka-* under the agent voice, and all the potentive and stative *ma-/na-* under the patient voice. It must be mentioned that there are a few instances that *ma-/na-* indicate the agent voice instead, e.g., *maligo* ‘take a shower’. According to Schachter and Otones

Table 4: Potentive voice, aspect, mood paradigm for *kain* ‘eat’ based on Himmelmann (2005b).

	Agent voice	Patient voice <in>	Patient voice –an	Patient voice i-/ipag-
Non-realis/Perfective	maka-kain	ma-kain	ma-kain-an	ma-i-kain
Non-realis/Imperfective	maka-ka~kain	ma-ka~kain	ma-ka~kain-an	ma-i-ka~kain
Realis Perfective	naka-kain	na-kain	na-kain-an	na-i-kain
Realis Imperfective	naka-ka~kain	na-ka~kain	na-ka~kain-an	na-i-ka~kain

(1972), these different uses of *ma-/na-* were previously distinguishable through vowel length – shorter length for the agent voice, and longer for the patient voice; but this distinction is no longer observed nowadays.

The different voice affixes can also be combined with verbs that have causative stem-forming prefix *pa-* which indicates that an entity causes another entity to act/move (Schachter and Reid 2008). The causer is treated as an agent, and the causee as a non-agent. In (24a), the *pa-* morpheme is combined with the patient voice, while in (24b), it is combined with the agent voice.

- (24) a. Pa-kain-in mo na ‘yung zebra
 CAU-eat-PV.INF 2.SG.NSBJ already SBJ zebra
 ‘Make the zebra eat/Feed the zebra.’ (C4)
- b. Mag-pa~pa-kulo daw muna (a)ko-ng tubig
 AV-CAU~IPFV-boil first 1.SG.SUBJ-LIN water
 ‘I will first make water boil/I will first boil water.’ (C5)

Appendix C: Hopper and Thompson’s (1980) parameters of transitivity can be found on Table 5

Table 5: Hopper and Thompson’s (1980) parameters of transitivity.

	High transitivity	Low transitivity
1 Participants	2 or more participants, agent and non-agent	1 participant
2 Kinesis	action	non-action
3 Aspect	telic	atelic
4 Punctuality	punctual	non-punctual
5 Volitionality	volitional	non-volitional
6 Affirmation	affirmative	negative
7 Mode	realis	irrealis
8 Agency	agent high in potency	agent low in potency
9 Affectedness of the non-agent	non-agent totally affected	non-agent not affected
10 Individuation of the non-agent	non-agent highly individuated	non-agent non-individuated

We use the term *non-agent* for Hopper and Thompson’s ‘O’/‘object’

References

- Adams, Karen L. & Alexis Manaster-Ramer. 1988. Some questions of topic/focus choice in Tagalog. *Oceanic Linguistics* 27(1/2). 79–101.

- Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122(3). 192–203.
- Aldridge, Edith. 2002. Nominalization and wh-movement in Seediq and Tagalog. *Language and Linguistics* 3(2). 393–442.
- Ambridge, Ben, Evan Kidd, Caroline F. Rowland & Anna L. Theakston. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42(2). 239–273.
- Amora, Kathleen Kay, Rowena Garcia & Natalia Gagarina. 2020. Tagalog adaptation of the Multilingual Assessment Instrument for Narratives: History, process and preliminary results. *ZAS Papers in Linguistics* 64. 221–233.
- Baayen, R. Harald, Anna Endresen, Laura A. Janda, Anastasia Makarova & Tore Nesset. 2013. Making choices in Russian: Pros and cons of statistical methods for rival forms. *Russian Linguistics* 37(3). 253–291.
- Bautista, Maria Lourdes S. 1983. Word order in Tagalog child language. *Philippine Journal of Linguistics* 14(1). 25–44.
- Billings, Loren. 2005. Ordering clitics and postverbal R-expressions in Tagalog: A unified analysis? In Andrew Carnie, Heidi Harley & Sheila Ann Dooley (eds.), *Verb first: On the syntax of verb-initial languages*, 303–339. Amsterdam & Philadelphia: John Benjamins.
- Bock, Kathryn, Helga Loebell & Randal Morey. 1992. From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review* 99(1). 150–171.
- Bresnan, Joan. 2007. A few lessons from typology. *Linguistic Typology* 11(1). 297–306.
- Brown, Roger. 1973. *A first language: The early stages*. London: George Allen & Unwin.
- Budwig, Nancy. 1990. The linguistic marking of nonprototypical agency: An exploration into children's use of passives. *Linguistics* 28(6). 1221–1252.
- Buenaventura Naylor, Paz. 1975. Topic, focus, and emphasis in the Tagalog verbal clause. *Oceanic Linguistics* 14(1). 12–79.
- Casillas, Marisa, Elika Bergelson, Anne S. Warlaumont, Alejandrina Cristia, Melanie Soderstrom, Mark VanDam & Han Sloetjes. 2017. A new workflow for semi-automatized annotations: Tests with long-form naturalistic recordings of children's language environments. In Francisco Lacerda (ed.), *18th annual conference of the international speech communication association (Interspeech 2017)*, 2098–2102. New York: Curran Associates.
- Chen, Feinian, Luoman Bao, Rachel M. Shattuck, Judith B. Borja & Socorro Gultiano. 2017. Implications of changes in family structure and composition for the psychological well-being of Filipino women in middle and later years. *Research on Aging* 39(2). 275–299.
- Chen, Victoria & Bradley McDonnell. 2019. Western Austronesian voice. *Annual Review of Linguistics* 5. 173–195.
- Cooreman, Ann, Barbara A. Fox & T. Givón. 1984. The discourse definition of ergativity. *Studies in Language* 8(1). 1–34.
- Croft, William. 2003. *Typology and universals*, 2nd edn. Cambridge: Cambridge University Press.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2021. *Ethnologue: Languages of the world*, 24th edn. Dallas, TX: SIL International.
- ELAN (Version 5.9) [Computer software]. 2020. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. <https://archive.mpi.nl/tla/elan>.
- Ferreira, Fernanda. 1994. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language* 33(6). 715–736.
- Foley, William A. 1998. Symmetrical voice systems and precategoriality in Philippine languages. In *Paper presented at the workshop on voice and grammatical functions in Austronesian languages*. University of Queensland, 30 June–3 July.

- Galang, Rosita G. 1982. Acquisition of Tagalog verb morphology: Linguistic and cognitive factors. *Philippine Journal of Linguistics* 13(2). 1–15.
- Garcia, Rowena. 2016. Thematic role assignment and word order preferences in the L1 acquisition of Tagalog. In *Paper presented at the IDEALAB Winter school*. University of Trento, 20 January.
- Garcia, Rowena. 2021. *Tagalog corpus*. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. <https://hdl.handle.net/1839/b6f082fe-05ca-4bf2-a1db-bf8ce63e58c8>.
- Garcia, Rowena, Jeruen E. Dery, Jens Roeser & Barbara Höhle. 2018. Word order preferences of Tagalog-speaking adults and children. *First Language* 38(6). 617–640.
- Garcia, Rowena & Evan Kidd. 2020. The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development* 16(4). 399–425.
- Garcia, Rowena, Gabriela Garrido Rodriguez & Evan Kidd. 2021. Developmental effects in the online use of morphosyntactic cues in sentence processing: Evidence from Tagalog. *Cognition* 216. 104859.
- Garcia, Rowena, Jens Roeser & Barbara Höhle. 2019. Thematic role assignment in the L1 acquisition of Tagalog: Use of word order and morphosyntactic markers. *Language Acquisition* 26(3). 235–261.
- Gleitman, Lila R., Kimberly Cassidy, Rebecca Nappa, Anna Papafragou & John C. Trueswell. 2005. Hard words. *Language Learning and Development* 1(1). 23–64.
- Gries, Stefan Th. 2020. On classification trees and random forests in corpus linguistics: Some words of caution and suggestions for improvement. *Corpus Linguistics and Linguistic Theory* 16(3). 617–647.
- Guilfoyle, Eithne, Henrietta Hung & Lisa Travis. 1992. SPEC of IP and SPEC of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10. 375–414.
- Hellwig, Birgit & Dagmar Jung. 2020. Child-directed language – and how it informs the documentation and description of the adult language. *Language Documentation & Conservation* 14. 188–214.
- Himmelmann, Nikolaus P. 1987. *Morphosyntax und Morphologie: Die Ausrichtungsaffixe im Tagalog*. Munich: Wilhelm Fink.
- Himmelmann, Nikolaus P. 2005a. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. London: Routledge.
- Himmelmann, Nikolaus P. 2005b. Tagalog. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 350–376. London: Routledge.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Linguistic Society of America* 56(2). 251–299.
- Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational & Graphical Statistics* 15(3). 651–674.
- Katagiri, Masumi. 2005. Voice, ergativity, transitivity in Tagalog and other Philippine languages: A typological perspective. In I Wayan Arka & Malcolm Ross (eds.), *The many faces of Austronesian voice systems*. Canberra: Pacific Linguistics.
- Kempen, Gerard & Karin Harbusch. 2004. A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary approaches to language production*, 173–182. Berlin & New York: Mouton de Gruyter.
- Kidd, Evan & Seamus Donnelly. 2020. Individual differences in first language acquisition. *Annual Review of Linguistics* 6. 319–340.

- Kroeger, Paul R. 1993a. Another look at subjecthood in Tagalog. *Philippine Journal of Linguistics* 24(2). 1–16.
- Kroeger, Paul R. 1993b. *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
- Labov, William. 1989. The child as linguistic historian. *Language. Variation and Change* 1(1). 85–94.
- LaPolla, Randy J. 2014. Constituent structure in a Tagalog text. *Language and Linguistics* 15(6). 761–774.
- Latrouite, Anja. 2011. *Voice and case in Tagalog: The coding of prominence and orientation*. Düsseldorf: Heinrich-Heine University dissertation.
- Law, Paul. 2016. The syntax of Tagalog relative clauses. *Linguistics* 54(4). 717–768.
- Levshina, Natalia. 2016. Why we need a token-based typology: A case study of analytic and lexical causatives in fifteen European languages. *Folia Linguistica* 50(2). 507–542.
- Levshina, Natalia. 2021. Conditional inference trees and random forests. In Magali Paquot & Stefan Th. Gries (eds.), *Practical handbook of corpus linguistics*, 611–643. New York: Springer.
- Lidz, Jeffrey & Annie Gagliardi. 2015. How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics* 1(1). 333–353.
- Lieven, Elena. 2016. Usage-based approaches to language development: Where do we go from here? *Language and Cognition* 8(3). 346–368.
- Lieven, Elena, Heike Behrens, Jennifer Speares & Michael Tomasello. 2003. Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30(2). 333–370.
- MacLachlan, Anna & Masanori Nakamura. 1997. Case-checking and specificity in Tagalog. *The Linguistic Review* 14. 307–333.
- MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum.
- Mahboob, Ahmar & Priscilla Cruz. 2013. English and mother-tongue-based multilingual education: Language attitudes in the Philippines. *Asian Journal of English Language Studies* 1. 1–19.
- Manueli, Maria Christina S. 2010. A minimalist account of scrambling and word order in Tagalog: Testing its grammaticality. *Philippine Social Sciences Review* 62(2). 161–195.
- Marzan, Jocelyn Christina B. 2013. *Spoken language patterns of selected Filipino toddlers and pre-school children*. Diliman: University of the Philippines dissertation.
- Marzan, Jocelyn Christina B., Maryelle Ellaine L. Dy, Mary Seanne A. Angeles, Jessica Francesca P. Arnaldo, Ma. Bettina Clare N. Camacho, Thea Beatriz G. Cruz, Ancellie L. Espino, Johannes C. Ibardaloza, Doris Renee V. Leaño, Chef Stan L. Macaraeg, Anna Christen D. L. Magbanua & Grace Camille S. Tan. 2014. *Correctness of verb inflections coding aspect and tense used by five 14 to 44-month-old Filipino children: An exploratory descriptive longitudinal study*. Manila: University of the Philippines Bachelor's thesis.
- Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Press.
- McDonnell, Bradley. 2016. *Symmetrical voice constructions in Besemah: A usage-based approach*. Santa Barbara: University of California dissertation.
- Minkoff, Seth. 2000. Animacy hierarchies and sentence processing. In Andrew Carnie & Eithne Guilfoyle (eds.), *The syntax of verb initial languages*, 201–212. Oxford: Oxford University Press.
- Nolasco, Ricardo Ma. 2005. What ergativity in Philippine languages really means. In *Proceedings of the Taiwan-Japan joint workshop on Austronesian languages*, 215–238. Taipei: National Taiwan University.

- Nolasco, Ricardo Ma. & Maureen Joy Saclot. 2005. M-and S-transitivity in Philippine type languages. In *Proceedings of the international course and conference on role and reference grammar*, 26–30. Taipei: Academia Sinica.
- Philippine Statistics Authority. 2003. National Capital Region: Close to 10 million persons (Results from the 2000 census of population and housing, NSO). <http://www.psa.gov.ph/population-and-housing/node/428> (accessed 2 March 2021).
- Philippine Statistics Authority. 2016. Highlights on household population, number of households, and average household size of the Philippines (2015 census of population). <https://psa.gov.ph/population-and-housing/node/69728> (accessed 25 February 2021).
- Phillips, Juliet R. 1973. Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development* 44(1). 182–185.
- Pine, Julian M. 1994. The language of primary caregivers. In Clare Gallaway & Brian J. Richards (eds.), *Input and interaction in language acquisition*, 15–37. Cambridge: Cambridge University Press.
- Pizarro-Guevara, Jed Sam & Matthew Wagers. 2020. The predictive value of Tagalog voice morphology in filler-gap dependency formation. *Frontiers in Psychology* 11(517). <https://doi.org/10.3389/fpsyg.2020.00517>.
- R Core Team. 2016. *R: A language and environment for statistical computing (Version 4.0.3) [computer program]*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36(4). 565–599.
- Reid, Lawrence A. & Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. *Language and Linguistics* 5(2). 433–490.
- Riesberg, Sonja. 2014. *Symmetrical voice and linking in Western Austronesian languages* (Pacific Linguistics 646). Berlin & Boston: De Gruyter Mouton.
- Riesberg, Sonja, Maria Bardaji i Farré, Kurt Malcher & Nikolaus P. Himmelmann. 2021. Predicting voice choice in symmetrical voice languages: All the things that do not work in Totoli. *Studies in Language* 46(2). 453–516.
- Riesberg, Sonja, Kurt Malcher & Nikolaus P. Himmelmann. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3). 523–561.
- Riesberg, Sonja & Beatrice Primus. 2015. Agent prominence in symmetrical voice languages. *STUF – Language Typology and Universals* 68(4). 551–564.
- Sabbagh, Joseph. 2009. Existential sentences in Tagalog. *Natural Language & Linguistic Theory* 27(4). 675–719.
- Sabbagh, Joseph. 2011. Adjectival passives and the structure of VP in Tagalog. *Lingua* 121(8). 1424–1452.
- Saffran, Jenny R. & Natasha Z. Kirkham. 2018. Infant statistical learning. *Annual Review of Psychology* 69. 181–203.
- Schachter, Paul. 2015. Tagalog. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax: Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science vol. 42/3), 1658–1676. Berlin & Boston: De Gruyter Mouton.
- Schachter, Paul & Fe T. Otnes. 1972. *Tagalog reference grammar*. Berkeley, CA: University of California Press.
- Schachter, Paul & Lawrence A. Reid. 2008. Tagalog. In Bernard Comrie (ed.), *The world's major languages*, 2nd edn. 833–855. London: Routledge.

- Segalowitz, Norman S. & Rosita G. Galang. 1978. Agent-patient word-order preference in the acquisition of Tagalog. *Journal of Child Language* 5(1). 47–64.
- Siewierska, Anna. 1993. On the interplay of factors in the determination of word order. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research*. Berlin & New York: Walter de Gruyter.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.
- Strobl, Carolin, James Malley & Gerhard Tutz. 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14(4). 323–348.
- Szmrecsanyi, Benedikt. 2017. Variationist sociolinguistics and corpus-based variationist linguistics: Overlap and cross-pollination potential. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62(4). 685–701.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.
- Tanaka, Mikihiro N., Holly P. Branigan, Janet F. McLean & Martin J. Pickering. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65(3). 318–330.
- Tanaka, Nozomi. 2016. *An asymmetry in the acquisition of Tagalog relative clauses*. Hawaii: University of Hawaii at Manoa dissertation.
- Tanaka, Nozomi, William O'Grady, Kamil Deen, Cha-Eun Kim, Ryoko Hattori, Ivan Paul M. Bondoc & Jennifer U. Soriano. 2014. Acquisition of Tagalog focus system. In *Paper presented at the 16th annual international conference of the Japanese Society for Language Sciences (JSLS2014)*. Saitama, Japan: Bunkyo University.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based account of language acquisition*. Cambridge, MA: Harvard University Press.
- Van Valin, Robert D. Jr & Randy J. LaPolla. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.
- Weiner, E. Judith & William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19(1). 29–58.