Cognition 216 (2021) 104859

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Developmental effects in the online use of morphosyntactic cues in sentence processing: Evidence from Tagalog

Rowena Garcia [a,*], Gabriela Garrido Rodriguez [a,b,c], Evan Kidd [a,c,d]

[a] Language Development Department, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[b] School of Languages and Linguistics, The University of Melbourne, Australia
[c] ARC Centre of Excellence for the Dynamics of Language, Canberra, Australia
[d] Research School of Psychology, The Australian National University, Canberra, Australia

ARTICLE INFO

Keywords:
Online sentence comprehension
Thematic roles
Morphosyntax
Language acquisition
Tagalog

ABSTRACT

Children must necessarily process their input in order to learn it, yet the architecture of the developing parsing system and how it interfaces with acquisition is unclear. In the current paper we report experimental and corpus data investigating adult and children's use of morphosyntactic cues for making incremental online predictions of thematic roles in Tagalog, a verb-initial symmetrical voice language of the Philippines. In Study 1, Tagalog-speaking adults completed a visual world eye-tracking experiment in which they viewed pictures of causative actions that were described by transitive sentences manipulated for voice and word order. The pattern of results showed that adults process agent and patient voice differently, predicting the upcoming noun in the patient voice but not in the agent voice, consistent with the observation of a patient voice preference in adult sentence production. In Study 2, our analysis of a corpus of child-directed speech showed that children heard more patient voice- than agent voice-marked verbs. In Study 3, 5-, 7-, and 9-year-old children completed a similar eye-tracking task as used in Study 1. The overall pattern of results suggested that, like the adults in Study 1, children process agent and patient voice differently in a manner that reflects the input distributions, with children developing towards the adult state across early childhood. The results are most consistent with theoretical accounts that identify a key role for input distributions in acquisition and language processing.

1. Introduction

In order to comprehend spoken language, listeners must rapidly process a transient linear signal in the moment. There is a general consensus that this process proceeds in an *incremental* and *interactive* manner (Altmann & Mirkovic, 2009; Altmann & Steedman, 1988; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Thothathiri, Asaro, Hsu, & Novick, 2018), making use of both top-down and bottom-up information to gradually build a semantic representation of an event. Accordingly, decades of psycholinguistic research with adults have shown that speakers make use of multiple cues during sentence processing, such as word order, case marking, and semantics, to make a series of moment-by-moment predictions as a sentence unfolds (e.g., Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003). Thus, the mature speaker deftly negotiates the problem of unpacking the dense information compressed in language by seizing upon reliable cues to structure and meaning.

Languages differ in how these cues are weighted (Bates & MacWhinney, 1989), and so the child language learner is faced with the seemingly difficult task of acquiring cues and their differential weightings in order to implement them online. Importantly, since many classic and current models of sentence processing assume that the parser makes use of distributional information that is necessarily acquired across development (e.g., Bates & MacWhinney, 1989; Chang, Dell, & Bock, 2006; Hale, 2001; Levy, 2008; MacDonald, 2013; MacDonald, Pearlmutter, & Seidenberg, 1994), empirical investigations of sentence processing that have an explicit developmental focus are needed.

In the current paper, we report on the development of online sentence interpretation in Tagalog, an Austronesian language spoken in the Philippines, with a focus on thematic role assignment (i.e., '*who* is doing *what* to *whom*'). The language has unique typological properties that allow us to test how variation in distributional patterns influence the development of parsing strategies. Moreover, since Tagalog is verb-initial and has a flexible ordering of arguments, our research crucially

* Corresponding author at: Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands.
E-mail address: rowena.garcia@mpi.nl (R. Garcia).

https://doi.org/10.1016/j.cognition.2021.104859
Received 6 October 2020; Received in revised form 8 July 2021; Accepted 12 July 2021
Available online 27 July 2021

widens the scope of psycholinguistic research, which has notably focused on a narrow and skewed set of the world's languages (Jaeger & Norcliffe, 2009; Lieven & Stoll, 2010).

1.1. Online parsing in development

Modern work on children's online parsing began some 20 years ago with the development of child-friendly eye-tracking systems (e.g., Trueswell, Sekerina, Hill, & Logrip, 1999) and an increased use of electroencephalography (EEG) in particularly difficult-to-test pre-school participants (e.g., Hahne, Eckstein, & Friederici, 2004). While this work has gone some way towards removing artificial barriers between the fields of adult language processing and language acquisition, we still lack explicit theoretical accounts of how children's emerging linguistic knowledge is implemented online. Explaining the development of the parsing system is a core aim of the field: children must necessarily parse their input in order to learn it (Fodor, 1998), and explaining developmental phenomena is a key desideratum in evaluating the explanatory scope of theories of adult parsing.

Studies on varying linguistic structures, age groups, and languages have shown a broad set of possibilities to explain children's acquisition and online implementation of linguistic knowledge. Following traditional themes in language acquisition research, the range of theoretical possibilities lie across a continuum, from those that primarily emphasize the use of innately-guided abstract representations and principles that are operational early in development (e.g., Phillips & Ehrenhofer, 2015), to those that emphasize input-driven representations that significantly change across development (e.g., Bates & MacWhinney, 1989; Chang et al., 2006; Seidenberg & MacDonald, 1999), with some occupying a location in the middle ground (e.g., Messenger & Fisher, 2018; Özge, Kuntay, & Snedeker, 2019).

Theoretical approaches that draw heavily on children's (and adults') experience with language to make parsing decisions include *experience-based* models of language processing, which assume that listeners incrementally use incoming information and its corresponding distribution to calculate the most probable continuation of the sentence (connectionist and constraint-based models: e.g., Chang et al., 2006; MacDonald, 2013; expectation-based models: e.g., Hale, 2001; Levy, 2008). For example, in expectation-based models, structures are built by choosing grammatical rules with high probabilities (essentially continuations that are frequent in the input) and the ease of processing is related to how closely the generated structure matches the structure that is needed to understand the meaning of the sentence (Levy, 2008). While successful in explaining adult processing data (e.g., reading times), these models do not explain syntax acquisition because the grammatical rules are assumed to be adult-like and unchanging; they are a given and assumed component of the theories. On the other hand, connectionist models using error-based learning can acquire different syntactic rules for different languages as well as encode the frequency of these rules (e.g., Chang, 2009 for English/Japanese; Chang, Baumann, Pappert, & Fitz, 2015 for German; Janciauskas & Chang, 2018 for Korean/English; and Tsoukala, Broersma, van den Bosch, & Frank, 2021 for Spanish/English). As such, the approach aligns with emergentist models of acquisition that place high importance on input-driven acquisition of knowledge (Ambridge, Kidd, Rowland, & Theakston, 2015; Bates & MacWhinney, 1989; MacWhinney, 1999).

Early abstraction accounts of acquisition claim that children have abstract representation of structures early in development. However, these accounts differ in how much they consider the role that input plays (if any). Specifically, it has been claimed that by the age of two or three, children have already abstracted the mapping between grammatical structures and thematic roles (Gertner, Fisher, & Eisengart, 2006), and can employ this knowledge for online thematic role assignment (Özge et al., 2019). For Gertner et al. (2006, p. 690), children do not need exposure to the input to learn these mappings, as they have predispositions for linking these semantic and structural abstractions. Similarly, Lidz, Gleitman, and Gleitman (2003), p. 163) claim that children learn argument structure patterns from universal mappings between the lexicon and syntax, and only the adults in their experiments considered the distribution of syntax-semantic relations in their ambient language. Moreover, Phillips and Ehrenhofer (2015, pp. 437-438) claim that, in general, sparse input does not lead to a delayed acquisition (and thus processing) of linguistic phenomena, since children can rely on abstract grammatical knowledge during parsing.

In contrast to the above accounts, other early abstraction accounts take into consideration the varying cue availability and reliability in the input to explain the cross-linguistic differences found in developmental studies of online processing (Özge et al., 2019). For example, the low frequency of passives in English is assumed to affect children's ability to correctly represent the passive form (Messenger & Fisher, 2018). Studies on dative alternation show that even if children acquire adult-like representations early on (Snedeker, 2013; Thothathiri & Snedeker, 2008), input distribution may have an effect on online processing, depending on the task. Thothathiri and Snedeker argue that children's experience with particular verbs can influence their performance in tasks that require the connection between the lexical verb and the structure, but not in tasks which only require abstract knowledge, such as preferential looking studies involving novel verbs.

Thus, the points of theoretical interest are whether: (i) children deduce abstract grammatical knowledge, which is early developing *and* adult-like in scope, and implement this online, with or without the influence of distributional usage patterns, or (ii) children gradually induce language-specific grammatical knowledge from the distributions in their input and use this knowledge to parse their input. At one end of the continuum, an experience-driven parsing system predicts that children's and adults' parsing decisions are intimately tied to distributional properties of the input, with connectionist approaches predicting that frequency-driven grammatical knowledge will result in observable developmental differences. At the other end of the continuum, grammatically abstract parsing predicts early adult-like parsing that is adult-like once a form has been acquired, and which may or may not be influenced by frequency. Studying online parsing *across* development and into adulthood is thus of key importance.

In spite of the claim of a direct relationship between children's grammatical knowledge and its online implementation (Lidz, 2021; Snedeker, 2013, pp. 208, 211–212[1]), studies have shown that children and adults do show differences in parsing, most notably in terms of speed of processing and their ability to recover from incorrect initial parsing decisions (Choi & Trueswell, 2010; Dick, Wulfeck, Krupa-Kwiatkowski, & Bates, 2004; Montgomery, Evans, Gillam, Sergeev, & Finney, 2015; Kidd, Stewart, & Serratrice, 2011; Trueswell et al., 1999). Incorrect parsing decisions based on what might initially be innately-guided linguistic generalizations (Gertner et al., 2006; Lidz et al., 2003) provide important evidence for a need to generate new hypotheses based on language-specific cues (Pozzan & Trueswell, 2015). This leads to the hypothesis that early appearing cues in an utterance (e.g., voice inflection on the verb in a verb-initial language) are more easily acquired than late appearing cues (e.g., voice inflection on the verb in an SOV language), because early-appearing cues allow an earlier retreat from 'garden-path' effects (see Huang, Zheng, Meng, & Snedeker, 2013).

1.2. Past research on children's parsing

There is an emerging literature on children's online parsing that bears upon these theoretical claims (e.g., Abbot-Smith, Chang, Rowland, Ferguson, & Pine, 2017; Huang, Zheng, Meng, & Snedeker, 2013; Özge

[1] Snedeker (2013) writes: "abstract representations play a dominant role in on-line comprehension in young children" and "children who have just turned three have abstract grammatical representations which they employ during online comprehension."

et al., 2019; Schipke, Friederici, & Oberecker, 2011; Schipke, Knoll, Friederici, & Oberecker, 2012; Strotseva-Feinschmidt, Schipke, Gunter, Brauer, & Friederici, 2019; Yang, Chan, Chang, & Kidd, 2020; Zhou & Ma, 2018). Here, we concentrate on those studies that investigated the cues that children use to assign core thematic roles (i.e., agent, patient/theme). The majority of the world's languages make use of one or two prominent linguistic means to mark thematic roles, using either word order or case marking on nouns.[2] Thus, a core issue in developmental studies of sentence processing has been how children simultaneously acquire and use these cues (Bates & MacWhinney, 1989). The current evidence is unclear as to how this process proceeds, with notable differences reported across languages.

In an offline picture selection task, Dittmar, Abbot-Smith, Lieven, and Tomasello (2008) reported that German-speaking children relied on word order to interpret transitive Noun-Verb-Noun (NVN) sentences until the age of seven, when they could already use case marking. This is despite German using case to mark thematic roles. Online studies using eye-tracking and EEG have shown similar results. Using the visual world eye-tracking paradigm, Kröger, Münster, and Knoeferle (2017) reported that 5-year-old German-speaking children do not use case markers to predict the second noun in transitive NVN sentences. Instead, upon hearing the first noun and the verb, they looked more towards the patient regardless of whether the sentence was agent- or patient-initial, suggesting that they interpreted the first noun as the agent regardless of its case marking. Schipke et al. (2012) showed converging evidence using EEG in a study that compared 6-year-old German-speaking children and adults. Overall, the data from German suggest that children do not use case as a cue to thematic role assignment until after 6 years, which Friederici and colleagues argue to become possible following the maturation of neural structures supporting grammatical processing (Friederici, 2011; see also Skeide, Brauer, & Friederici, 2016. The suggestion is that, rather than relying on abstract knowledge of the German case system, preschool-aged children rely on a word order strategy by interpreting the first noun as the agent of the sentence, regardless of case.

In contrast to the German data, eye-tracking studies in children acquiring Turkish or Mandarin have shown more adult-like processing abilities, as indicated by their online use of morphosyntactic markers to predict upcoming referents, earlier in development (Huang et al., 2013; Özge et al., 2019; Zhou & Ma, 2018). In Mandarin, Zhou and Ma (2018) reported that 5-year-olds used the co-verbs *ba* (indicating that the following argument is the patient) and *bei* (indicating that the following argument is the agent) markers for thematic role assignment. In their study, transitive constructions with a dropped argument were used, so that the sentences were *ba/bei*-initial. They found that 5-year-olds already show predictive gaze patterns similar to adults: the children looked at the target picture immediately after hearing the marker and the following argument. In Turkish, Özge et al. (2019) showed that 4-year-old Turkish-speaking children appear capable of using nominative or accusative case marking on the first noun to predict the upcoming second noun, even without the help of verb information (in NNV sentences). For example, upon hearing an accusatively-marked *rabbit*, children directed more looks to the *fox* compared to the *carrot*, showing that they correctly interpreted the rabbit as the patient of the action and expected an agent of the action (*fox* is more likely to be the agent of the action *eat* compared to the *carrot*) to be mentioned next. The authors argued that the results support early abstraction accounts, as children were able to assign abstract thematic roles independent of the verb (Özge et al., 2019, p. 169).

However, it is important to point out that the evidence for the suggestion that young children can use purely morphological cues to predict thematic roles online, and thus rely solely on abstract morphosyntactic knowledge to make early parsing predictions, is still equivocal. The best evidence comes from Experiment 2 in Özge et al. (2019), who reported that children looked significantly more to agents before the second noun in NNV sentences when the first noun was accusative marked. The effects for the children, however, are difficult to interpret when compared to the adult data. Children seemed to show faster processing compared to adults earlier in the sentence, but *slower* processing once they hear the second noun. Moreover, in all past studies on Turkish and German, the use of morphosyntactic information was determined through prediction of the second argument; both word order (first noun) and morphosyntactic information (case marker) were available by the disambiguating region. A similar problem arises in the case of Mandarin: in Zhou and Ma (2018), either the *ba* and *bei* marker and the following argument were required to determine the target picture. These findings therefore cannot show us whether morphosyntactic information is used when the word order cue is not available, however reliable that cue might be. Thus, it remains an open question as to whether children can use morphosyntactic information alone to predict the thematic role of the upcoming first noun, and whether their ability to do so is mostly experience-dependent or indicative of the use of abstract syntactic generalizations implemented online. In this paper, we investigated this phenomenon in Tagalog, a verb-initial language that does not formally use word order for assigning thematic roles; but instead, uses highly reliable morphosyntactic markers.

1.3. Tagalog

Tagalog is an Austronesian language from the Philippines, with more than 23 million native speakers (Simons & Fennig, 2018). It is considered a *symmetrical voice* language; that is, it has multiple basic transitive forms, and has a corresponding marker for each voice alternation (Foley, 1998; Himmelmann, 2005a; Riesberg, 2014). It is verb-initial and the marking on the verb assigns the thematic roles of the arguments (Himmelmann, 2005b).[3] The agent voice inflection -*um*- assigns the subject, the *ang*-phrase, an agent role (Ex. 1, 3); while the patient voice inflection -*in*- assigns the *ang*-phrase a patient role (Ex. 2, 4). The order of the post-verbal arguments is relatively free, so both agent-initial (Ex. 2, 3) and patient-initial sentences (Ex. 1, 4) are possible in both voices.

(1)	H<**um**>ahabol	ng	manok	**ang**	lalaki
	<AV>[a]chase	NSBJ	chicken	SBJ	man
	'The man is chasing a chicken.'				
(2)	H<**in**>ahabol	ng	manok	**ang**	lalaki
	<PV>chase	NSBJ	chicken	SBJ	man
	'The/A chicken is chasing the man.'				
(3)	H<**um**>ahabol	**ang**	lalaki	ng	manok
	<AV>chase	SBJ	man	NSBJ	chicken
	'The man is chasing a chicken.'				
(4)	H<**in**>ahabol	**ang**	lalaki	ng	manok
	<PV>chase	SBJ	man	NSBJ	chicken
	'The/A chicken is chasing the man.'				

[a] AV refers to the agent voice, PV to the patient voice, SBJ to subject, and NSBJ to non-subject.

Thus, when interpreting a basic transitive sentence like those in (1)–(4), a Tagalog speaker must, procedurally, link the voice marking to the noun markers to determine argument relations. This mapping in sentences (1)–(4) is categorical, and thus abstract knowledge of each marker theoretically enables the prediction of upcoming arguments, as would unambiguous case markers in a language like Turkish. It is important to note that, unlike most languages studied in

[2] Note that, while these types of languages are most prominently studied in Psycholinguistics, there is a substantial number of languages that do not follow this pattern. For instance, many polysynthetic languages have flexible word order and no or very little case marking on nouns, instead marking core arguments in the verb phrase (e.g., Evans, 2003; Nordlinger, 2011).

[3] A subject-initial structure is also grammatical but it is considered formal and it is mostly found in texts (Schachter & Otanes, 1972).

Psycholinguistics, the grammatical subject in Tagalog does not default to the agent. On the contrary, analyses of a written corpus(Cooreman, Fox, & Givón, 1984) and child-directed speech (CDS; Garcia, Roeser, & Höhle, 2019) suggest that the patient voice agent-initial construction (2) is the most frequently used structure (thus mapping the prominent *ang*-argument to the patient). Thus, analyses of naturalistic data suggest that adult Tagalog speakers most commonly construe transitive events from the perspective of the patient, and identify it as the prominent syntactic argument. In the current paper, we present data from a new corpus of CDS to better estimate these frequencies.

Only a few studies have been conducted on the online processing of Tagalog. Using the visual world paradigm, where adult participants saw displays containing pictures of a possible agent, possible patient, and a distractor (e.g., a frog, a fly, and a computer printer), Sauppe (2016) found that adult native speakers directed their looks to the agent (e.g., the frog) after hearing the verb (e.g., '*eat*') regardless of its voice-marking. This finding shows that the voice-marking alone did not affect prediction of the upcoming arguments; rather, Sauppe found that participants' eye-movements were influenced by voice in concert with the first noun and its marker. That the participants did not use the voice marking early but instead looked to the agent may be a consequence of Sauppe's method. Specifically, since participants were required to establish the most probable agent-patient relations between referents (i.e., a frog is more likely to eat a fly than vice-versa), participants likely needed to first establish an event representation of the argument hierarchy before language-specific mechanisms could come into play.

Using eye-tracking and a sentence-picture matching task, Garcia, Roeser, and Höhle (2020) presented adults and children with two pictures of reversible actions between animals (e.g., *monkey bites cat, cat bites monkey*) while they heard verb-initial sentences crossed for voice (i.e., agent versus patient) and the order of arguments (i.e., agent-first versus patient-first). Results showed that adults immediately looked at the target picture upon hearing the voice-marked verb and noun marker + noun combination, suggesting rapid use of this early information. In contrast, 5- and 7-year-old children showed a different pattern of results. After hearing the verb and the first noun marker and its noun, 7-year-olds preferred to look at the picture which showed the first noun as the agent compared to the picture where it was shown as the patient, indicating a word order strategy. In the temporal adverb region following the first noun, the 5-year-olds did not yet show a preference for any picture. However, after hearing the 2nd noun, both 5-year-olds and 7-year-olds looked at the target picture, although for patient-initial sentences, they showed better performance in the patient voice than in the agent voice. Similar processing preferences in favor of the patient voice were reported by Garcia et al. (2019) in a separate sample of 5- and 7-year-old children using self-paced listening. Additionally, consistent with its higher frequency in the input, offline studies (i.e., picture selection or verification) of the acquisition of the Tagalog voice system suggest a patient voice advantage (Galang, 1982; Segalowitz & Galang, 1978; and offline accuracy data reported in Garcia et al., 2019; 2020).

In summary, the previous online studies on Tagalog show that adults incrementally process the morphosyntactic markers for thematic role assignment, while children's online use of the markers seem to be limited to the patient voice and is still not as efficient as the adults', even at age 7-years. However, since past studies analyzed the first noun marker and the first noun together, we do not yet know the influence of the morphosyntactic markers separate from that of the first noun in online thematic role assignment. In other words, it is unclear whether children use the morphosyntactic information on the verb to *predict* the upcoming first noun, and whether their ability to use morphosyntactic marking is experience-dependent, based on an abstract understanding of the noun markers, or a combination of the two. In the current research, we used a version of the visual world eye-tracking paradigm that allowed us to investigate Tagalog-speaking adult's and children's use of purely morphosyntactic cues to predict thematic roles.

1.4. Current research

We report three studies that investigated the use of morphosyntactic markers in online thematic role assignment by Tagalog-speaking adults and children aged 5–9 years. Study 1 reports on a visual world eye-tracking study of Tagalog-speaking adults' online processing of simple transitive sentences manipulated for voice and argument order, thereby determining the adult end-state that children are developing towards. In particular, we focus on whether Tagalog-speaking adults can rely solely on the voice-marking on the verb and on the first noun marker to predict the identity of the first noun (e.g., agent or patient), which would constitute the most unambiguous evidence that the verb and noun markers are used as cues to thematic role assignment. Study 2 presents an analysis of child-directed speech, enabling a more accurate formulation of hypotheses for children's online use of cues, given the assumptions of accounts that draw heavily on the input distribution and those that do not. In Study 3 we explicitly examined these competing accounts by testing 5-, 7-, and 9-year-old children on a slightly modified version of Study 1.

2. Study 1: Tagalog-speaking adults' processing of transitive sentences

In Study 1, we used the visual world paradigm to investigate Tagalog-speaking adults' processing of transitive sentences manipulated for voice (agent vs. patient) and word order (agent-first vs. patient-first), in a fully crossed within-participants design. Our focus was on whether Tagalog-speaking adults use the verbal voice-markers and the noun markers to rapidly predict upcoming referents online. While Tagalog is a symmetrical voice language, voice and argument order are not equal in their distributions, which experience-based models predict will influence online parsing decisions (Chang et al., 2006; Hale, 2001; Levy, 2008; MacDonald, 2013). Two points are particularly relevant here: (i) adult native speakers of Tagalog overwhelmingly prefer to use the patient voice to describe transitive events (Bondoc, O'Grady, Deen, & Tanaka, 2018; Sauppe, Norcliffe, Konopka, Van Valin Jr., & Levinson, 2013; Tanaka, 2016), and (ii) sentence completion studies using materials similar to ours (i.e., transitive events involving two animate referents) show that, whereas the patient voice word order follows a relatively fixed Verb-Agent-Patient (VAP) ordering (> 90%), word order in the agent voice is equally distributed across VAP and VPA orders (both at approximately 50%) (Garcia, Dery, Roeser, & Höhle, 2018).

The different accounts of parsing make contrasting predictions for online processing in adults. Assuming that adults have acquired the grammatical rules of the language, a parser that relies predominantly on abstract grammatical rules to generate structures in real time should show similar behaviors for these simple transitive structures (Phillips, 2013; Phillips & Lewis, 2013), such that they will be able to predict the upcoming first noun in both voices. Expectation-based theories assume that adults have perfect knowledge of the rules of the language and all possible structures are evaluated in parallel using the probabilistic information (Hale, 2001; Levy, 2008). These theories predict a speed advantage for the more frequent patient voice structures, but there should not be a large difference in the ability to distinguish VAP and VPA across both voices. Finally, connectionist accounts assume that each of these structures is learned, and because the voice information occurs before nouns, it can have a large impact on the structural representations (Chang, 2009). Specifically, if processing differs by voice, that suggests that there is not some abstract VAP and VPA structures common to both voices, but rather that there are *distinct* structures for each voice (agent-VAP, agent-VPA, patient-VAP, patient-VPA). This account predicts a larger advantage for the processing of patient voice structures.

2.1. Method

2.1.1. Participants

We recruited 32 adult native speakers of Tagalog from a university in Metropolitan Manila (mean age: 21, range: 18–34, females: 9). All participants had normal or corrected-to-normal vision, and provided written consent. None of them had a history of speech or language delay, nor neurologic or psychiatric disorders. All of the participants reported proficiency in English which is typical for native Tagalog speakers in Metropolitan Manila (Amora, Garcia, & Gagarina, 2020).[4] A few participants reported being proficient in other Philippine languages (e.g., Cebuano, $n = 2$; Rinconada, $n = 1$), and in other foreign languages (e.g., Farsi, n = 1; Japanese, n = 1). Ethics approval was granted by the Ethiek Commissie Sociale Wetenschappen (ECSW-2018-041, Amendment ECSW2017–3001-474).

2.1.2. Materials

Sixteen Tagalog transitive verbs (e.g., *hila* 'pull') were used in the stimuli sentences. The verbs were chosen because they can be inflected with the agent and patient voice infixes, both of their arguments can be animate, and they are easy to visually depict.

Each verb was combined with two animal pairs (from a pool of 8 common animals) resulting in 32 verb-animal pair combinations. Each verb-animal pair combination appeared in all of the four experimental conditions, namely agent voice agent-initial, agent voice patient-initial, patient voice agent-initial, and patient voice patient-initial, resulting in a total of 128 experimental sentences (see Table 1 for sample experimental items and Appendix A for a complete list of the sentences). Crucially, this means that, unlike in past studies (e.g., Özge et al., 2019; Sauppe, 2016), noun phrase (NP) semantics could not be used as a cue to a referent's thematic role, and thus any evidence for predictive processing prior to noun onset will be attributable to the formal features of the morphosyntactic markers. The number of times that an animal was used as the agent and patient in a sentence was counterbalanced across the experiment. In each experimental sentence, an adjective was placed after the first noun marker to prolong the period before the first noun is heard, thus providing us time to observe how the voice-marking on the verb and the first noun marker are used to predict the upcoming first noun, and giving the participants more time for prediction. Two native speakers of Tagalog (i.e., the first author and a graduate student of Linguistics) checked the combination of the adjectives and verb-animal pairs such that no adjective was more likely to occur with any animal nor any thematic role in each item. Additionally, a temporal adverb was added after the verb to prolong the time before the first noun marker was mentioned. The experimental sentences were recorded by a native Tagalog speaker (the first author) in an audio booth. The sentences were recorded with a normal speaking rate on Audacity software program (version 2.0.5; Audacity Team, 2015). The average duration of each sentence region during a trial is in Table 2.

Thirty-two fillers and three practice items were also prepared and audio-recorded. These items did not contain verbs, and were descriptions such as 'There is a spoon on the big table' and 'There are many coconuts in the store today.' All stimulus pictures were created by a professional artist (see Fig. 1 for an example). Each picture was around 635 × 315 pixels in size. The direction of action in the experimental picture in each trial and throughout the experiment was counterbalanced. The mirror image of each experimental picture was used for two of the experimental conditions, and the original image was used for the other two. The animals in each picture and throughout the experiment were of similar size.

The 128 experimental sentences and their corresponding pictures were distributed into 4 lists using a Latin square design. Each list contained 32 experimental trials (8 items per condition) with the 16 verbs appearing twice, but with each verb-animal pair combination appearing only once. Each list was divided into two blocks, with each verb being presented only once in each block. The experimental items were interspersed with the fillers, so no two experimental items followed each other. Furthermore, the experimental items were pseudo-randomized such that the same voice or order condition was not presented more than three times in a row. Each participant was randomly assigned to one list.

2.1.3. Procedure

Participants were individually tested in a quiet room at the university. The experiment was presented using SMI Experiment Center (version 3.7, SensoMotoric Instruments GmbH, 2017) on a 17-in. laptop with a resolution of 1024 × 768 pixels. Below the laptop screen, an SMI RED-mobile eye-tracker (120 Hz sampling rate) was placed to record the participants' eye movements. The auditory stimuli were presented through closed headphones. The experimenter sat next to the participant only before and after the main eye-tracking experiment. During the eye-tracking experiment, the experimenter sat away from the participant and monitored the flow of the experiment through a separate screen.

Before the experiment, individual pictures of the animals and actions that would appear in the experiment were shown on the computer screen. In the action pictures, two boys perform the action instead of two animals. Four pictures were presented at a time. These pictures were presented to the adult participants, to make the procedure consistent with that of Study 3, where the experimenter first had to check whether the children were familiar with the animals and verbs that would be used in the experiment. The experimenter then proceeded to a five-point calibration and four-point validation of the eye-tracker.

After the calibration and validation phase, three practice trials that were similar to the fillers were presented. Participants were told that they would see pictures and hear sentences corresponding to these pictures, and they should listen carefully because at the end of the experiment they will be asked questions about what they saw and heard. At the beginning of each trial, participants had to look at a fixation circle (presented at the top center of the screen) for at least 500 ms before the experiment program would present the next picture. This gaze-contingent presentation of the visual stimuli was done so the participants' gaze would not land incidentally on the agent or the patient of the action even before the picture was presented. The stimulus picture was presented in the center of the screen with a grey background. After 1000 ms of practice/filler picture presentation, the corresponding audio-recorded sentence was played. The picture remained on the screen throughout the audio presentation, and for around 1000 ms after the end of the sentence. Each practice or filler item was 5000 ms long. After the three practice items, two type-written *yes/no* questions were presented on the screen one after the other: *May aso ka bang nakita?* "Did you see a dog?" and another question that could only be answered by listening to the auditory stimuli *Mayroon bang ibinibenta sa iyong mga nakita?* "Was there something for sale among the things you saw?" Participants had to verbally answer these questions.

After the practice trials, the main experiment was presented. Because the experimental pictures were more complex than the practice/filler pictures, the audio-recorded sentence was played only after 1500 ms of picture presentation. The picture remained on the screen throughout the audio presentation, and for around 1000 ms after the end of the sentence. Each experimental trial lasted for 7000 ms. We presented the picture before and throughout the audio presentation to reduce task demands. Moreover, we chose to present only one picture at a time instead of a picture selection task, as the latter may force the participants to consider alternative interpretations that they may not have otherwise

[4] English is an official language in the Philippines, and the country has a long history of using English as a medium of instruction in schools (Tupas & Lorente, 2014), so Tagalog native speakers usually have some proficiency in English. Aside from Tagalog, other Philippine-languages are commonly used in households in Metro Manila (Mahboob & Cruz, 2013; Philippine Statistics Authority, 2003).

Table 1
Sample experimental sentences for the verb *kagat* 'bite' in Study 1. The vertical lines indicate the sentence regions: verb + temporal adverb, 1st noun marker + adjective, 1st noun, and 2nd noun marker + second noun.

(1) Agent voice agent-initial	K<um>akagat	noong Martes	ang	masipag	na	baka	ng	unggoy
	<AV>bite	last Tuesday	SBJ	diligent	LIN	cow	NSBJ	monkey
	'The diligent cow was biting a monkey last Tuesday.'							
(2) Agent voice patient-initial	K<um>akagat	noong Martes	ng	masipag	na	unggoy	ang	baka
	<AV>bite	last Tuesday	NSBJ	diligent	LIN	monkey	SBJ	cow
	'The cow was biting a diligent monkey last Tuesday.'							
(3) Patient voice agent-initial	K<in>akagat	noong Martes	ng	masipag	na	baka	ang	unggoy
	<PV>bite	last Tuesday	NSBJ	diligent	LIN	cow	SBJ	monkey
	'The/A diligent cow was biting the monkey last Tuesday.'							
(4) Patient voice patient-initial	K<in>akagat	noong Martes	ang	masipag	na	unggoy	ng	baka
	<PV>bite	last Tuesday	SBJ	diligent	LIN	monkey	NSBJ	cow
	'The/A cow was biting the diligent monkey last Tuesday.'							

thought about (Zuckerman, Pinto, Koutamanis, & van Spijk, 2016). The current task is more similar to picture description tasks (e.g., Griffin & Bock, 2000; Sauppe et al., 2013), as we assumed that once the participants saw the picture during the preview, they built an expectation about the sentence that would be used to describe the visual stimulus. Consequently, this method minimized the chances of observing a word order (or first-noun-as-agent) strategy. We suggest that this strategy is more likely to be observed in tasks where a semantic representation of an event is being constructed as part of the experimental task because participants must simultaneously build an event representation, which is most commonly done from the perspective of the highest order argument (i.e., the agent, MacWhinney, 1999), while mapping that emerging representation onto incoming language. In the current study, the participants did not have to build their own event construal, as it was already made available by the picture shown before the audio-recording of the sentence was played.

Table 2
Average duration (ms) of each sentence region in experimental trials of Study 1.

	Verb + Temporal adverb	1st noun marker + adjective	1st noun	2nd noun marker +2nd noun
1. Agent voice agent-initial	2103	893	752	768
2. Agent voice patient-initial	2054	908	740	736
3. Patient voice agent-initial	2071	919	754	735
4. Patient voice patient-initial	2071	886	745	768

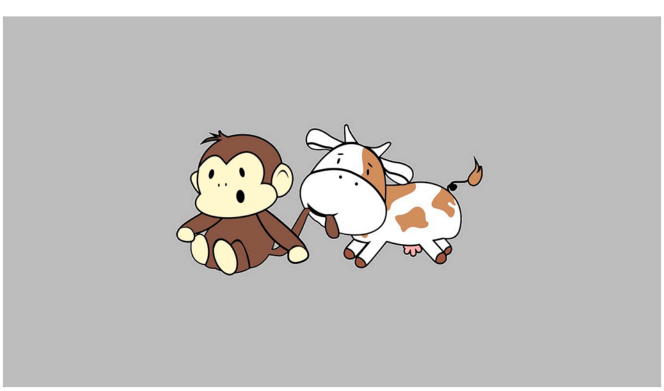

Fig. 1. Sample visual stimulus for experimental sentences with the verb *kagat* 'bite' in Study 1.

After the first block, participants were told that half of the experiment was already finished. After that, the second block was presented. The experiment finished with five *yes/no* questions (e.g., *Mayroon bang regalo para kay Tatay?* "Was there a gift for Father?") but the answers were no longer recorded, as these comprehension questions were included only as a cover task for participants to pay attention to the stimuli. The whole session lasted for approximately 15 min.

2.1.4. Data analysis

Three areas of interest (agent, patient, and action region) were manually drawn for each experimental item using the SMI BeGaze software (version 3.7, SensoMotoric Instruments GmbH, 2017). The part of the image that showed the interaction between the two animals (e.g., the mouth of the cow biting the tail of the monkey in Fig. 1) was considered as the action region. We isolated this region because it was not clear whether fixations to this area were looks to the agent or to the patient. There was no overlap among the three areas of interest. BeGaze was also used to export the raw eye-tracking data, which include information about the stimulus and gaze behavior (e.g., fixations, saccades and blinks as implemented by the manufacturer's algorithm) with respect to the areas of interest (e.g., fixations could land on the agent, patient, action region or white space), for each recording time point (i.e., every 8.33 ms). All data pre-processing and statistical analyses were performed in R software (version 3.6.2, R Core Team, 2016). Data points showing a saccade in one eye and a fixation in the other were removed, as well as data points showing different areas of interest for the left and right eye. Intervening blinks in blocks of fixations towards the same region were also turned into fixations to the same region. Moreover, trials with fixation transitions longer than 600 ms were removed (1.9%

of the data).

There is no standard way to analyse visual-world eye-tracking data, the major problem being how the statistical models (e.g., ANOVAs, generalized mixed-effects models) deal with time as a variable. One common way to do this is to break the sentence into segments or time windows (e.g., words or constituents) and analyse each one separately, but for several reasons this approach is suboptimal. Firstly, it places artificial boundaries on processing events, which likely traverse different elements in the linguistic signal (or alternatively, occur in smaller sections within a pre-defined segment). Secondly, past studies using this analytic approach have often failed to correct for multiple comparisons in their repeated analyses of the eye-movement behavior, which increases the possibility of making a Type I error.

To overcome these problems, we used a non-parametric permutation analysis (Chan, Yang, Chang, & Kidd, 2018; Good, 2005; Groppe, Urbach, & Kutas, 2011; Maris & Oostenveld, 2007; Maris, 2012). This technique builds a sampling distribution (i.e., the permutation distribution) by resampling the observed data, and is ideal for our purposes because it identifies processing events in the eye-tracking record in a data-driven manner, determining the time in the eye-tracking record where looks to the agent diverge across word order conditions for each voice type. Thus, if there are more fixations to the agent in the agent-initial condition than in the patient-initial condition immediately after the first noun marker is encountered, but before the first noun is mentioned, we can conclude that the participants used the voice-marking on the verb and the noun marker to predict the upcoming first noun. The data was analyzed following a series of steps (for a detailed explanation and implementation of the test, see Chan et al., 2018). In the first step, *linear regression* models were conducted to evaluate the effects of word order (categorical independent variable coded as agent-initial = 0.5, patient-initial = −0.5) on fixations to the agent (vs. patient) for every 8.33 ms time bin (following the eye-tracker's sampling at 120 Hz). The models were conducted separately for each verbal voice (i.e., agent voice and patient voice). Thus, we modelled the proportion of fixations to the agent as a function of word order in agent voice and in patient voice. The proportion of fixations to the agent was calculated by dividing the number of fixations to the agent by the total number of fixations to the agent and to the patient. Fixations towards the action region were not included in the analysis because it could not be determined whether these were supposed to be looks to the agent or to the patient. The regressions provided a list of time bins with significant *p*-values (i.e., black bars above −0.01 represent significant values ($p < .05$), see Fig. 2). In the second step, significant adjacent time bins were *clustered*, under the assumption that they likely constitute a single processing event. In the final step, a *permutation distribution* was created by randomly permuting the order and voice labels of the clusters in order to fit a regression model on this randomized data. The procedure was repeated 1000 times. The outcome of this procedure provides a distribution of sum *t*-values for each cluster, which shows the likelihood that a cluster had occurred by chance.

2.2. Results

Fig. 2 shows the mean proportion of looks to the agent against time for the agent-initial and patient-initial conditions in each voice (agent voice and patient voice). In the *patient voice* sentences, adults looked more to the agent in the agent-initial condition compared to the patient-initial condition before the first noun was mentioned (from 4281 ms until 4656 ms [375 ms duration], sum $t = 143$, $p < .001$). There were more looks to the agent in the patient-initial condition than in the agent-initial condition following the second noun marker and during the second noun (5781 ms–5997 ms [216 ms duration], sum $t = 60$, $p < .001$; 6005 ms–6072 ms [67 ms duration], sum $t = 18$; $p = .002$; 6455 ms–6522 ms [67 ms duration], sum $t = 19$; $p = .002$). In the *agent voice* sentences, no anticipatory looks to the agent for the agent-initial condition were observed. The eye-tracking data showed more looks to the agent in the patient-initial compared to the agent-initial condition during the second noun (5764 ms–6305 ms [541 ms duration], sum $t = 180$; $p < .001$; 6355 ms–6438 ms [83 ms duration], sum $t = 27$; $p < .001$). Additionally, for a short cluster during the verb region, there were more looks to the agent in the patient-initial condition than in the agent-initial condition (1715 ms–1782 ms [67 ms duration], sum $t = 18$, $p = .02$).

2.3. Discussion

In Study 1, we found that adult speakers of Tagalog rapidly predicted the first noun from the morphosyntactic markers in the more frequent patient voice, but did not do so for the less frequent agent voice. This is difficult to explain within a parsing mechanism that strongly adheres to grammatical principles and relies on purely abstract form-meaning mappings between voice and the noun markers (e.g., Phillips, 2013; Phillips & Lewis, 2013). The abstract mapping account predicts no voice difference, but in contrast to the patient voice results, we found no evidence of predictive looks to the first noun in the agent voice. Indeed, we did not observe any significant looks to the first noun at all.

Expectation-based approaches (Hale, 2001; Levy, 2008) assume that comprehenders have access to the same grammatical machinery but that processing speed is moderated by frequency, predicting a similar pattern of effects across voice type but with a speeded patient voice advantage. While the results showed a patient voice advantage, the lack of difference in looks to the first noun referent in the agent voice is suggestive of a greater difference than a simple lag in processing speed, and instead suggesting that different processes were taking place in each voice. This is more consistent with theories where structural rules can vary depending on preceding linguistic elements (Chang, 2009). On this approach, the early prediction of the first noun in the patient voice is attributable to both the higher frequency of this voice type and the relative stability of argument order in transitive sentences containing that voice.

Two additional results deserve mention. First, we found a short significant cluster in the beginning of the verb region. This result could be due to factors that were not related to our manipulated linguistic variables. However, we suggest that it is also fairly meaningless. Importantly, this early cluster did not seem to affect looking patterns downstream, since there was no difference in the looks to the agent between conditions for around 3900 ms after the initial cluster appeared. Second, unlike Sauppe (2016), we found no evidence that adult speakers of Tagalog direct their looks to the possible agent upon hearing the verb, independent of voice-marking. Sauppe attributed this finding to the importance of agents in building a mental representation of the event. The difference between the two studies is likely methodological. Namely, whereas the participants in Sauppe's study necessarily had to construct an event representation because the stimuli did not depict an action, our stimuli depicted a simple and unambiguous event.

3. Study 2: Corpus study of Tagalog child-directed speech

In Study 1, we found evidence that Tagalog-speaking adults can use morphosyntactic markers to predict upcoming referents, but that this ability was crucially dependent on distributional evidence available in the language, as indexed by adult production patterns. Since a key focus of our research is how children develop parsing strategies, we needed to accurately categorize the distributional patterns in child-directed speech. Although an existing child language corpus of Tagalog exists (Marzan, 2013), it is relatively small in size. Thus, we collected a new corpus of Tagalog child-caregiver interaction in order to have a more accurate estimate of the relative frequency of voice and argument order before we conducted Study 3.

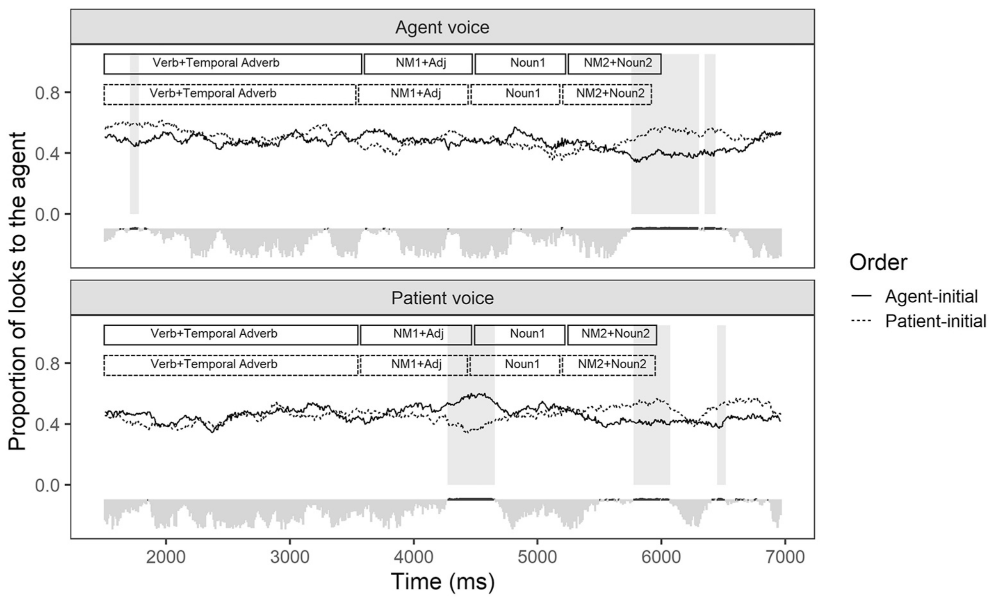

Fig. 2. Adults' average proportion of looks to the agent from verb onset until the end of the trial in Study 1. The sentence regions are indicated by the rectangles (NM1 refers to the first noun marker, Adj to adjective, NM2 to the second noun marker). The small grey/black bars around −0.01 indicate the *p* value from the linear regression for each time bin. Grey bars (below −0.01) indicate a *p* value greater than .05, while black bars (above −0.01) indicate a significant *p* value. The large grey shadings indicate the time bins which were found to be significant in the permutation analysis.

3.1. Method

3.1.1. Participants

Twenty Tagalog-speaking child-guardian pairs from Caloocan city in Metropolitan Manila participated. We recruited children aged from 2;0 to 4;0 years, in order to get a broadly representative sample of CDS across the early childhood years. Twelve of the guardians were the children's mothers, five were grandmothers, two were fathers, and one was the grandfather (mean age: 36, age range: 21–72). Based on a questionnaire, all guardians were native Tagalog speakers. Moreover, they used mostly Tagalog in their households. Eight guardians reported proficiency in other Philippine languages (Kapampangan, $n = 3$; Bikol, $n = 2$; Cebuano, $n = 1$; Ilocano, n = 1; Pangasinan, n = 1). Socio-economic status, as defined by education level, was mixed: three university graduates, eight had some university education, seven high school graduates, one had a few years of high school, and one had finished elementary school.

3.1.2. Materials

Dyads were provided with a range of materials that aimed to elicit conversation. These included a range of toys, such as a kitchen set, doctor set, dolls, animal figures, cars, furniture miniatures, magic slate, and blocks. Different pictures of causative and non-causative actions with varying animacy of the entities were also printed and compiled into three albums. The wordless picture storybook *Frog, Where are You?* (Mayer, 1969) was also included in the set. We used a video camera with a microphone to record the child-guardian interactions.

3.1.3. Procedure

Guardians were encouraged to play with their child using the materials provided by the researcher (first author, a native speaker of Tagalog). The researcher remained as unobtrusive as possible, and did not interact with the dyad during the session unless it was absolutely necessary. No other instruction was given, and each recording lasted 60 min.

The typicality of one-on-one caregiver-child interactions varies across cultures; and given the average household size of 4.4 persons (Philippine Statistics Authority, 2016) and prevalence of extended family members living together in the Philippines (Chen, Bao, Shattuck, Borja, & Gultiano, 2017), one-on-one interactions might not be as common as in countries with a smaller average household size. Thus, while we acknowledge that the sessions do not likely mimic the exact experience of the Tagalog-learning child, our methodological set-up fulfilled our aim to elicit significant amounts of CDS in a semi-naturalistic setting.

3.1.4. Data analysis

The video recordings were transcribed by two research assistants (Linguistics graduate and Speech Pathology graduate who are native speakers of Tagalog) on ELAN (Version 5.9) [Computer software] (2020). The transcription rules were loosely based on the minCHAT format (MacWhinney, 2000), following the DARCLE Annotation Scheme (Casillas et al., 2017). The transcripts were further annotated on ELAN for voice inflection on the verb, word order, and noun-marking, by one of the research assistants and by the first author. The annotations were exported to txt files, and the calculations and statistical analyses were conducted in R software (version 3.6.2, R Core Team, 2016).

The child-directed utterances included in the current analysis were only verb-initial sentences with highly causative verbs, similar to the stimuli sentences used in Studies 1 and 3. Causativity was judged based on Hopper and Thompson's (1980) criteria (e.g., volitionality of the agent and affectedness of the patient), e.g., *hiniwa* 'sliced' was counted but not *hinanap* 'looked for.' Questions, and sentences with modals were excluded. We checked the most frequent order of agents and patients, as well as the frequency of the voice markers. The patient voice in this study refers to Himmelmann (2005b) undergoer voices, which is an umbrella term used for voice markers that assign the *ang*-phrase a non-agent role (e.g., patient, beneficiary, instrument, and goal). In all of these voices, the agent is marked with a *ng*. In the word order analyses, we considered these non-agent arguments as patients, and we did not include utterances with the pronoun *kita* (I to you).

Table 3
Percentage distribution of voice-inflected and uninflected causative sentences in the child-directed speech in Study 2.

	Agent voice	Patient voice	Uninflected
All utterances	8	45	47
Utterances with only an agent or a patient	8	50	42
Utterances with both agent and patient	8	43	49

3.2. Results and discussion

In the 26,461 child-directed utterances, we found 9403 verb phrases. Out of these verb phrases, 31% (2935) passed our criteria (verb-initial, highly causative utterances). Around half of the verbs in these utterances were not inflected for voice (see Table 3). Fifty-five percent of the uninflected verbs appeared in imperative or hortative utterances. The use of the base form of verbs for agent voice verbs in imperatives or hortatives (e.g., *Kain ka* instead of *Kumain ka* "You eat") is acceptable in adult-directed speech (Galang, 1982). However, a significant number of utterances containing these uninflected verbs, which were commonly descriptive utterances or questions, were in baby-talk register, in which using uninflected verbs is common. In addition, 14% of the uninflected verbs were English nouns or verbs used in Tagalog frames, such as *check-up*, *kiss*, and *drawing*.[5] This code-switching is common in the modern use of Tagalog (Bautista, 2004; Thompson, 2003).

The voice-inflected verbs were mostly in the patient voice, regardless of whether the utterance had both the agent and the patient expressed, or if either the agent or a patient was ellided, or if there was no argument. Additionally, 82% of sentences with uninflected verbs (and at least an agent or a patient) should have been in the patient voice, based on the context and the marking on the noun(s) (i.e., the *ang*-phrase denoted the patient). Thus the CDS data are consistent with the general adult preference to use the patient voice. Within each voice type as well as for uninflected verbs, a dominance of agent-initial/agent-only structure was also observed (see Table 4).

The dominance of the patient voice compared to the agent voice in transitive sentences, coupled with the high frequency of agent-initial sentences, is consistent with Garcia et al.'s (2019) analysis of Marzan's (2013) corpus. However, there is a slight difference in the word order distributions to Garcia et al.'s (2018) sentence completion data used to motivate hypotheses in Study 1, which showed that participants produced agent-initial and patient-initial sentences equally in the agent voice. This discrepancy is due to the inclusion of sentences with pronouns in the corpus counts, while the experiments only included full noun phrases. Pronouns have a stricter order in Tagalog—they must occur immediately after the verb or the predicate (Billings, 2005), and because most pronouns were agents, we found a high frequency of agent-initial sentences regardless of voice. If we remove pronouns from the analysis (leaving 331 utterances; see Table 5) and keep sentences with both an agent and a patient, we see a distribution that more closely approximates the sentence completion results: we found that 56% (10/18) of agent voice sentences are agent-initial, whereas 75% (29/39) of patient voice sentences are agent-initial.

Table 4
Percentage distribution of sentences with agent-initial and patient-initial orders within each voice type in child-directed utterances with or without pronouns in Study 2.

	Utterances with only an agent or a patient	Utterances with both agent and patient
Agent voice		
Agent-initial	62	92
Patient-initial	38	8
Patient voice		
Agent-initial	72	88
Patient-initial	28	12
Uninflected		
Agent-initial	75	96
Patient-initial	25	4

Table 5
Percentage distribution of sentences with agent-initial and patient-initial orders within each voice type in child-directed utterances without pronouns in Study 2.

	Utterances with only an agent or a patient	Utterances with both agent and patient
Agent voice		
Agent-initial	26	56
Patient-initial	74	44
Patient voice		
Agent-initial	39	74
Patient-initial	61	26
Uninflected		
Agent-initial	18	80
Patient-initial	82	20

Overall, the results suggest that: (i) children acquiring Tagalog hear more verbs in the patient voice in VAP word order, (ii) there is a general tendency for transitive sentences to have agent-patient word order, but (iii) this tendency is weakened when sentences contain two lexical NPs, which allows freer ordering of NP constituents, most prominently in the agent voice. Note that since a predominant VAP order was found in both voices, the complexity of the voice system means that this results in the opposite distribution of noun markers. That is, in the patient voice, the order is *ng*-before-a*ng*, but in the agent voice, it is the opposite. Therefore, while word order provides a mostly reliable cue to thematic role assignment, children's mastery of the noun marking system, which perfectly predicts the thematic role, involves overcoming this variability. Children hear many more transitive sentences in the patient voice, leading to the possibility that input-driven acquisition and processing of the Tagalog noun marking favors the patient voice.

4. Study 3: Tagalog speaking children's processing of transitive sentences

The analysis of CDS allows us to generate predictions for children's processing of these structures. Accounts that do not consider a substantive role of the input in acquisition (Lidz et al., 2003; Phillips & Ehrenhofer, 2015) predict that, since children should acquire abstract knowledge of the Tagalog voice system and its relation to the noun markers early on, they should be equally adept at using the markers to predict the upcoming first noun. Since both general processing speed and speed of language processing improve across childhood (Dick et al., 2004; Kail & Salthouse, 1994), early abstraction accounts do not rule out the possibility that children's parsing would get faster with age. However, since children are argued to base parsing decisions on abstract form-meaning mappings, these accounts do not predict a processing asymmetry based on voice. Finally, we also point out that, since the relevant morphology—the verbal infix and the first noun marker—appear at the beginning of the sentence, accounts of processing difficulty based on cue location do not predict notable difficulty for acquisition or processing (Pozzan & Trueswell, 2015).

In contrast, experience-based accounts (Chang et al., 2006; Hale, 2001; Levy, 2008; MacDonald, 2013) and accounts of early abstraction that identify a role for children's language experience (Messenger & Fisher, 2018; Özge et al., 2019; Thothathiri & Snedeker, 2008) predict that children's online processing will be guided by the distributional properties of the input. Given the higher frequency of the patient to that of the agent voice,[6] the prediction is that, like in Study 1, children would

[5] Note that code-switched English verbs are not always uninflected in Tagalog; rather, the borrowed words are productively inflected, suggesting they are fully incorporated into the grammar.

[6] Even if many of the verbs in the corpus were uninflected, most of these sentences showed a mapping of the nouns consistent with that of the patient voice (*ng* marking for the agent and *ang* for the patient).

expect to hear a patient voice-marked verb more than an agent voice-marked verb. Where these accounts differ is in their explanation of *developmental* effects. Connectionist models that learn their syntactic representations predict that children's online processing across development will reflect changes in input-driven syntactic generalizations. For instance, the Chang et al. (2006) model develops a temporary first-noun-as-agent bias for English, which has also been observed in human babies (Abbot-Smith et al., 2017). On these accounts, the high frequency of the patient voice should also lead to children's better mastery of the mapping of the patient voice verb infix to the noun markers for assigning thematic role, compared to that of the agent voice, and crucially that there will be developmental differences in the efficiency with which this knowledge is implemented online. Past experimental research by Garcia and Kidd (2020) showed that, whereas 7-year-old Tagalog-speaking children regularly use noun markers correctly in both voices, 5-year-olds only did so around 50–60% of the time, and were less accurate in the agent than in the patient voice. If we take the incorrect use of noun markers in obligatory contexts as indicative of a developmental effect in children's knowledge, then the experience-based approach predicts developmental difference in online processing between 5-year-olds and older children.

Early abstraction accounts that allow frequency to guide parsing decisions predict that parsing should be adult-like relatively early in development (Snedeker, 2013, see footnote 1). Thus, while the approach predicts the voice asymmetry observed in adults, it does not predict significant developmental differences in processing once children are 3 years and older for basic transitive sentences (see Özge et al., 2019).[7]

We tested these predictions in Study 3. We gave 5-, 7-, and 9-year-old Tagalog-speaking children a slightly modified version of the visual-world eye-tracking task used in Study 1 to investigate whether children use the voice-marking on the verb and the first noun marker to predict the upcoming first noun, and how this changes across development.

4.1. Method

4.1.1. Participants

We tested a total of 154 children from Metropolitan Manila, Philippines: 54 5-year-olds (mean age: 5;7, age range: 5;0–6;0, girls: 22), 50 7-year-olds (mean age: 7;6, age range: 7;1–7;11, girls: 30), and 50 9-year-olds (mean age: 9;6, age range: 9;0–10;0, girls: 32). The 5-year-olds were Kindergarten or Grade 1 students, the 7-year-olds were Grade 2 or Grade 3 students, and the 9-year-olds were Grade 4 students from a public elementary school. We obtained written consent from the guardians for the children to participate.

A questionnaire completed by the children's guardians showed that all of the participants were dominant in Tagalog, and they came from Tagalog-speaking households. Furthermore, we screened all participants using a Tagalog vocabulary test to ensure that all were typically developing. Three participants were excluded (e.g., one 5-year-old, one 7-year-old, and one 9-year-old) as they showed speech and language errors before and after the main experiment, which were not exhibited by their peers (e.g., using uninflected verbs). The final sample consisted of 151 participants (53 5-year-olds; 49 7-year-olds; 49 9-year-olds).

4.1.2. Materials

The experimental sentences, fillers, and practice sentences, along with their corresponding pictures were the same as those used in Study 1, except that audio-recorded *who*-questions were added in this study. A sample question in the agent voice is *Sino ang k**um**akagat?* "Who is biting?" and in the patient voice: *Sino ang k**in**akagat?* "Who is being bitten?" (see Appendix B for a complete list of the questions). Comprehension questions were added to check if the children knew the difference between the agent voice and patient voice inflection on the verb. Each list included 16 comprehension questions (i.e., only half of the trials were followed by a question). For Lists 1 to 4, eight verbs were always paired with an agent voice question, and the other eight were always paired with a patient voice question. We also created four additional lists (Lists 5–8) to counterbalance the pairing of the verb and the voice of the question. Moreover, the pictures used in Lists 5–8 were mirror images of those used in Lists 1–4.

Similar to Study 1, the items were pseudo-randomized in each list such that no experimental item followed each other, and that the same voice or order condition was not presented more than three times in a row. Moreover, the same type of question (e.g., whether it was in agent voice or in patient voice) and questions where the answer was the same animal, were not presented more than three times in a row. Participants were randomly assigned to one list.

4.1.3. Procedure

Two participants were tested at a time, in separate rooms of the school. Half of the participants completed the eye-tracking experiment first (with the first author), followed by the vocabulary test with another experimenter (native speaker research assistant); while the other half of participants first completed the vocabulary test and then the eye-tracking study. The same eye-tracker set-up from Study 1 was used, except that in this study, the experimenter sat next to the participant during the whole session.

Before the experiment, individual pictures of the animals and actions that would appear in the experiment were presented on a computer screen (four pictures were presented at a time). Children were asked to point to the item which the experimenter labelled. If participants made a mistake, they were reminded to look carefully at the pictures and the four items of that specific trial were presented again. Once all of the pre-experimental items had been correctly identified, the experimenter proceeded to a five-point calibration and four-point validation of the eye-tracker.

After the calibration and validation phase, practice trials similar to the fillers were given. Participants were told that they would see pictures and hear sentences corresponding to those pictures, and sometimes the pictures could be followed by questions, which they had to verbally answer. Instead of a fixation circle, cartoon characters were used as a gaze trigger for every trial. Participants were told that they had to look at the cartoon characters to see the next picture. The presentation of practice items, experimental items and fillers was similar to that of Study 1. However, for two out of the three items in the practice phase, and for half of the experimental items, the audio-recorded sentence was followed by 1500 ms of silence and then by an audio-recorded *who*-question (e.g., *Sino ang nasa kweba?* 'Who is in the cave?' for the practice trial). The visual stimulus was presented throughout the trial. The participants were asked to give a verbal response before the gaze trigger for the next item was presented. The experimenter also repeated the question whenever the participant did not hear the audio-recorded question. The experimenter manually recorded the responses. The entire eye-tracking experiment was also audio-recorded to be able to review the manually recorded responses. After the first block of trials, recalibration and revalidation were done. The whole experimental session (eye-tracking study and vocabulary test) lasted approximately 30 min.

4.1.4. Data analysis

For the eye-tracking data, the same procedure as in Study 1 was used. Trials with more than 50% track loss were removed (0.5% of the data). A permutation analysis was conducted per age group to analyse the eye-tracking data. For the accuracy data, a logistic mixed-effects regression was fitted in R software (version 3.6.2, R Core Team, 2016). using the glmer function of the lme4 package (version 1.1–23, Bates, Mächler,

[7] Transitive sentences in the agent voice are not rare in Tagalog, as in, for instance, a full BE passive in English. Thus, on the early abstraction accounts, 5-year-old children should not differ from older children in their mappings from voice inflection to noun marker.

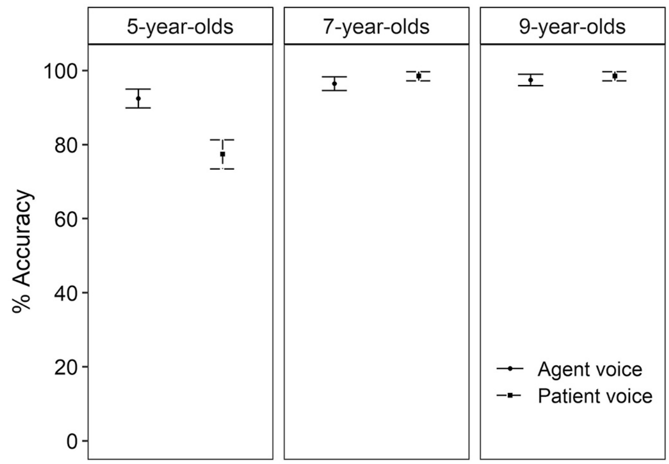

Fig. 3. Children's mean accuracy (%) and 95% confidence intervals in the agent and patient voice comprehension questions in Study 3.

Table 6
Model summary of the regression coefficients and variance components for the logistic mixed-effects model. Children's response accuracy was modelled as a function of age, voice of the question, and their interaction.

Fixed effects	β	SE	z value	p value
Intercept	4.91	0.43	11.39	<.001
Age (5:7)	−2.23	0.55	−4.07	<.001
Age (7:9)	−0.20	0.60	−0.33	.74
Voice (agent voice)	−0.97	0.41	−2.38	.02
Age (5:7)*Voice	1.48	0.51	2.90	<.01
Age (7:9)*Voice	−0.19	0.57	−0.34	.74
Random effects	**Variance**	**SD**		
Subject (intercept)	4.09	2.02		
Subject (voice slope)	3.09	1.76		
Item (intercept)	0.12	0.34		

Bolker, & Walker, 2015) to evaluate children's accuracy in the comprehension questions (coded as correct = 1, incorrect = 0) as a function of age (forward difference contrast coding: 5 vs. 7; 7 vs. 9), voice-marking of the question (sum contrast coding: agent voice vs. patient voice), and their interaction. The model that converged included random intercepts for subjects and items and by-subject slope adjustment for voice.

4.2. Results

4.2.1. Accuracy in the comprehension questions

The results showed above chance accuracy for all conditions (see Fig. 3). The logistic mixed-effects regression showed a significant main effect of age (5-year-olds vs. 7-year-olds) and voice-marking of the question, and an interaction of age (5-year-olds vs. 7-year-olds) and voice (see Table 6). Inspecting the interaction showed that 7-year-olds scored higher than the 5-year-olds but only in the patient voice ($\beta = -3.71$, $SE = 0.95$, $p = .001$), and no age difference was found in the agent voice ($\beta = -0.76$, $SE = 0.47$, $p = .58$).

4.2.2. Eye-tracking

The eye-tracking results presented here are only from the data of children who scored above 75% in the offline task.[8] This was to ensure that children already knew the difference between the two voice markers. Accordingly, data from 13 five-year-olds were excluded. We also excluded one item in List 1 due to an error in the experiment program (i.e., the same experimental item was included twice in the experiment, and the second presentation was excluded from the analysis).

The permutation analysis showed that in both the *agent* and *patient* voice sentences, 5-year-olds looked more to the agent in the agent-initial condition compared to the patient-initial condition only after the first noun had already been mentioned (see Fig. 4; agent voice: 5156 ms–5922 ms [766 ms duration], sum $t = 547$, $p < .001$; patient voice: 4673 ms–5863 ms [1190 ms duration], sum $t = 642$, $p < .001$), although the significant region began over 480 ms earlier in the patient voice. These looks persisted even while the second noun was being mentioned. It was only after hearing the second noun that the looks switched—the 5-year-olds looked more to the agent in the patient-initial condition (whose second noun was the agent) compared to the agent-initial condition (agent voice: 6255 ms–6996 ms [741 ms duration], sum $t = 352$, $p < .001$; patient voice: 6089 ms–6996 ms [907 ms duration], sum $t = 418$, $p < .001$).

In the *patient voice*, the 7-year-olds started to show more looks to the agent in the agent-initial condition compared to the patient-initial condition upon hearing the first noun marker (see Fig. 5; 4314 ms–4597 ms [283 ms duration], sum $t = 99$; $p < .001$). This gaze pattern persisted throughout the mention of the first noun and even partially into the second noun (4606 ms–5797 ms [1191 ms duration], sum $t = 594$, $p < .001$). After hearing the second noun, the looks switched (6005 ms-6746 [741 ms duration], sum $t = 337$, $p < .001$). In the *agent voice*, participants showed more looks to the agent in the agent-initial compared to the patient-initial condition only upon mention of the first noun (5172 ms–5730 ms [558 ms duration], sum $t = 302$, $p < .001$). The looks switched after hearing the second noun (6105 ms–6996 ms [891 ms duration], sum $t = 422$, $p < .001$).

Similar to the 7-year-olds, the 9-year-olds showed more looks to the agent in the agent-initial compared to the patient-initial condition in the *patient voice* upon hearing the first noun marker (see Fig. 6; 4339 ms–5505 ms [1166 ms duration], sum $t = 549$, $p < .001$). This result persisted until the first noun was mentioned. In the *agent voice*, there was a short significant cluster between 4489 ms and 4639 ms (150 ms duration, sum $t = 48$, $p < .001$; around the onset of the first noun). A longer cluster between the two orders was observed only after the first noun (5131 ms–5747 ms [616 ms duration], sum $t = 282$, $p < .001$), showing more looks to the agent in the agent-initial compared to the patient-initial condition. In both voices, there were more looks to the agent in the patient-initial condition than in the agent-initial condition upon hearing the second noun (agent voice: 5989 ms–6580 ms [591 ms duration], sum $t = 284$, $p < .001$; patient voice: 5806 ms–6738 ms [932 ms duration], sum $t = 506$, $p < .001$). Additionally, in the patient voice sentences, we also found more looks to the agent in the agent-initial compared to the patient-initial condition at the beginning of the verb region (1516 ms–1649 ms [133 ms duration], sum $t = 42$, $p < .001$).

4.3. Discussion

In Study 3, we examined the development of Tagalog-speaking children's ability to use morphosyntactic markers to make predictions about upcoming referents. The overall pattern of the data showed that, like the adults in Study 1, children in all age groups presented a patient voice advantage, which is consistent with the greater availability of the patient voice pattern in the input (Study 2). However, we also observed developmental differences in the data; notably, while the 5-year-old children did not show predictive looks to the first noun, the 7- and 9-year-old children did. Thus, like adults, Tagalog-speaking children process language incrementally, but their ability to use purely morphosyntactic cues to predict referents develops with age, tied to the distributional properties of the voice system. Therefore, the results are broadly consistent with experience-based accounts, which identify a key role for input distributions in acquisition and online parsing (Chang

[8] Including all of the children in the analysis shows similar results.

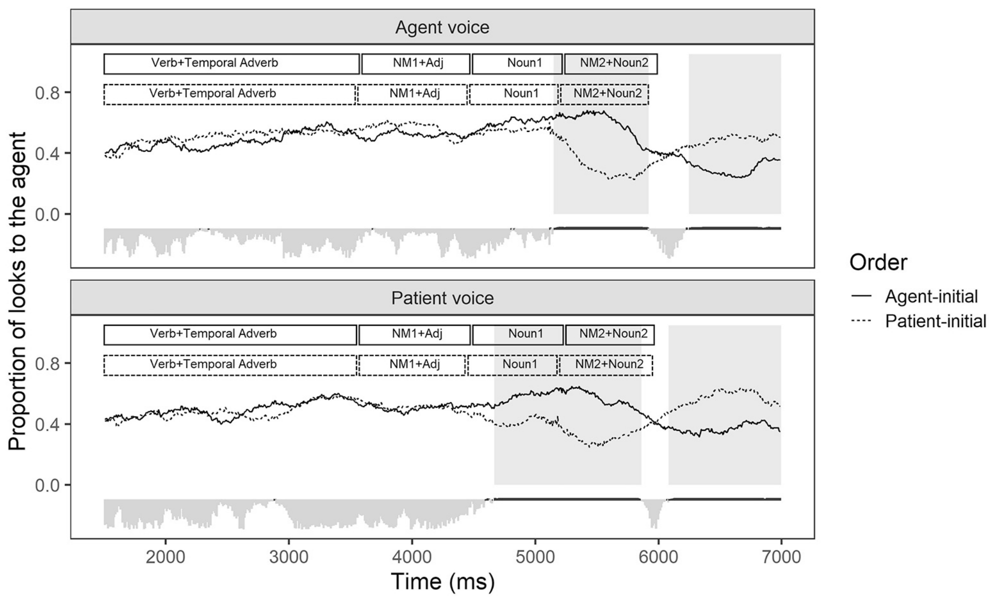

Fig. 4. Five-year-olds' average proportion of looks to the agent from verb onset until the end of the trial in Study 3. The sentence regions are indicated by the rectangles (NM1 refers to the first noun marker, Adj to adjective, NM2 to the second noun marker). The small grey/black bars around −0.01 indicate the *p* value from the linear regression for each time bin. Grey bars (below −0.01) indicate a *p* value greater than .05, while black bars (above −0.01) indicate a significant *p* value. The large grey shadings indicate the time bins which were found to be significant in the permutation analysis.

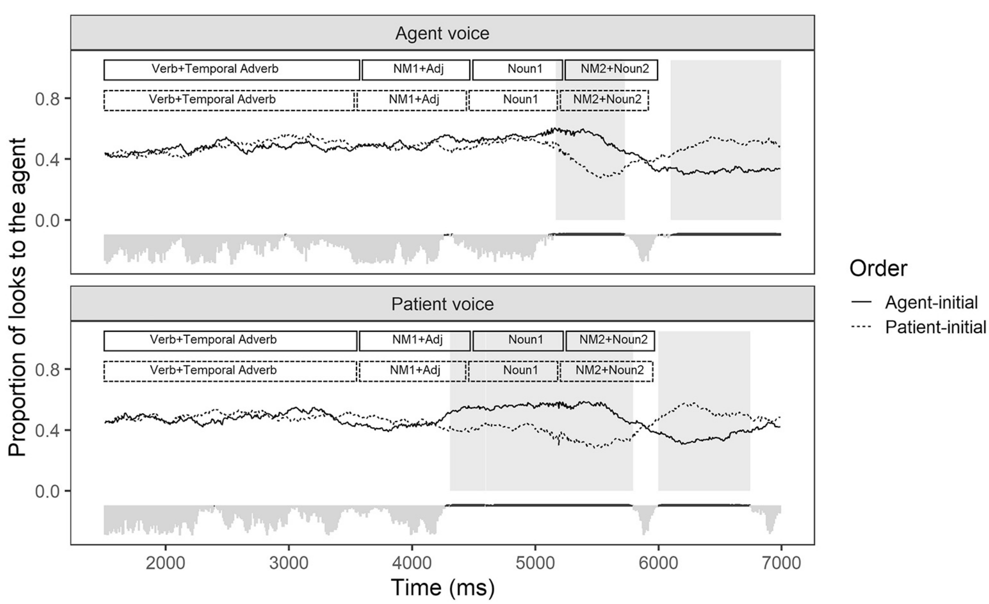

Fig. 5. Seven-year-olds' average proportion of looks to the agent from verb onset until the end of the trial in Study 3. The sentence regions are indicated by the rectangles (NM1 refers to the first noun marker, Adj to adjective, NM2 to the second noun marker). The small grey/black bars around −0.01 indicate the *p* value from the linear regression for each time bin. Grey bars (below −0.01) indicate a *p* value greater than .05, while black bars (above −0.01) indicate a significant *p* value. The large grey shadings indicate the time bins which were found to be significant in the permutation analysis.

et al., 2006; Hale, 2001; Levy, 2008; MacDonald, 2013), as well as early abstraction accounts that also consider children's language exposure (Messenger & Fisher, 2018; Özge et al., 2019; Thothathiri & Snedeker, 2008); but are inconsistent with accounts that do not consider the role of the input in early acquisition (Lidz et al., 2003; Phillips & Ehrenhofer, 2015).

Although the 5-year-old children did not show evidence of predictive use of the morphosyntactic markers in the patient voice, their pattern of eye-movements was qualitatively similar to that of the older children, such that their results suggest that they are developing towards becoming more efficient online parsers. In particular, the 5-year-olds demonstrated the same voice effect, as they identified the first noun in the patient voice more than 400 ms earlier than in the agent voice. Thus, their eye-movements are influenced by distributional information related to voice, but are not automated enough to enable prediction solely on the basis of the noun marker. We do note, however, that even the 5-year-olds converged on the first noun before its offset, and while it was still some 300 ms slower than their older peers, does suggest that they are rapidly making the correct parsing choices while they are hearing the first noun. The key empirical observation is that, by 5 years on-line processing in Tagalog is frequency sensitive but not yet adult-like. The key theoretical question concerns the underlying nature of this difference, which we address in the General Discussion.

Three other results deserve mention. Unlike the adults in Study 1, the 9-year-old children showed a short (150 ms) but significant cluster at the onset of the first noun in the agent voice condition. This could be interpreted as predictive processing in the agent voice which was not shown by the adults, supporting the claim of connectionist-based models that children have non-adult-like temporary biases that eventually disappear as their language skills develop (Chang et al., 2006). But if so, it is unclear as to why the children did not pursue this interpretation, as they did not look significantly more to the target until the first noun was mentioned. Because the significant window was short, we are hesitant to interpret the result as predictive processing. Secondly, there was a short cluster in the beginning of the verb region. Similar to the early cluster found in the adult eye-movement data, this result may be chance

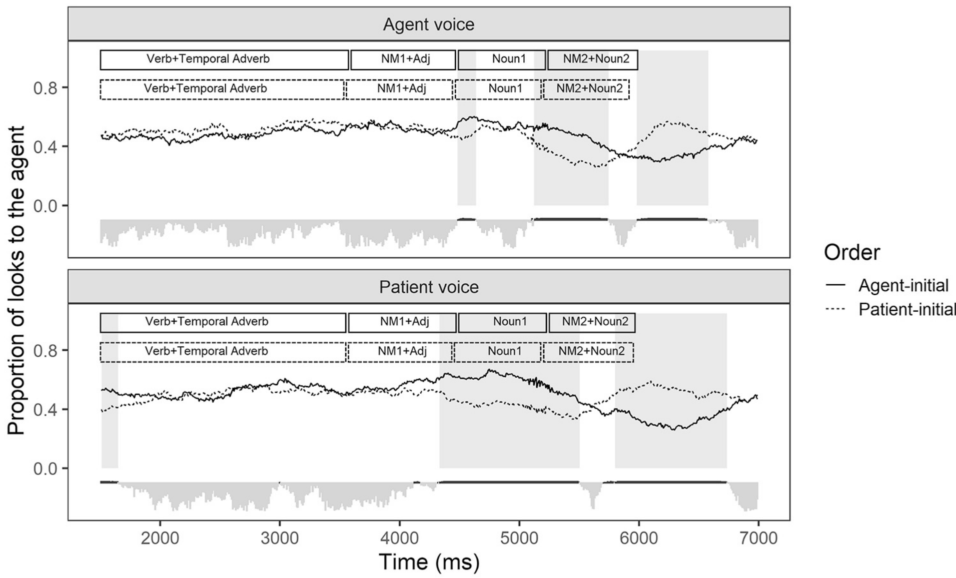

Fig. 6. Nine-year-olds' average proportion of looks to the agent from verb onset until the end of the trial in Study 3. The sentence regions are indicated by the rectangles (NM1 refers to the first noun marker, Adj to adjective, NM2 to the second noun marker). The small grey/black bars around −0.01 indicate the *p* value from the linear regression for each time bin. Grey bars (below −0.01) indicate a *p* value greater than .05, while black bars (above −0.01) indicate a significant *p* value. The large grey shadings indicate the time bins which were found to be significant in the permutation analysis.

deviation that was not related to our manipulated linguistic variables. More importantly, there was no difference in the looks to the agent between conditions for around 2700 ms after the initial cluster appeared.

Finally, the accuracy results also revealed an effect of age in the use of the patient voice marker. Older children showed better performance in answering comprehension questions in the patient voice compared to the 5-year-olds. However, there was no age difference observed in the agent voice. The agent voice advantage in the current study is probably due to the question word used—*sino* which usually refers to humans (Malicsi, 2013) who are more likely agents. This could explain why some 5-year-olds had difficulty with the patient voice, with a subset only answering the comprehension questions with agents regardless of the voice-marking in the comprehension question, resulting in low accuracy.

5. General discussion

Across three studies we investigated adults' and children's use of the morphosyntactic markers for assigning thematic roles and predicting the first argument in a verb-initial language. In doing so, we tested the claims of experience-based sentence processing accounts which predict that listeners incrementally use incoming information and its corresponding distribution to calculate the most probable continuation of a sentence (e.g., Chang et al., 2006; Hale, 2001; Levy, 2008; MacDonald, 2013), and early abstraction accounts which predict that children make early form-meaning abstractions that guide their syntactic choices online, including those that identify a role for distributional frequency information (Messenger & Fisher, 2018; Özge et al., 2019; Snedeker, 2013; Thothathiri & Snedeker, 2008), and those that do not (Lidz et al., 2003; Phillips & Ehrenhofer, 2015).

Overall, the data best support accounts that heavily consider the role of the input: despite the fact that there is a categorical mapping between voice and the thematic role denoted by noun markers, we found evidence for a patient voice preference in all four groups of participants we tested, which is consistent with the distributional differences consistently reported for Tagalog (Garcia et al., 2018; Sauppe et al., 2013; Study 2). Since the mapping is categorical, parsing accounts that privilege abstract grammatical knowledge without reference to frequency information predicts uniform use of the noun markers (e.g., Lidz et al., 2003; Phillips & Ehrenhofer, 2015), which we did not find.

Thus the data provide broad support for theoretical approaches that identify key roles for frequency information in the development and implementation of parsing preferences. Throughout this paper we have placed several different models under the umbrella of 'experience-based' approaches, and while these models typically converge in terms of predictions, they differ on other relevant dimensions. Notably, given our focus on the development of parsing, those models that simultaneously *acquire* language-specific grammatical knowledge through the business of processing their input, and hone parsing procedures based on further exposure to the input are preferable to models that operate with existing formal grammars. Thus, although experience-based models are difficult to distinguish based on general predictions concerning the influence of input distributions on parsing choices, explicitly learning-based models such as Chang et al. (2006, see also Dell & Chang, 2014) and MacDonald (2013), where acquisition and processing are explicitly input-driven, intertwined processes, are preferable to those experience-based models that annotate existing tagged corpora with usage frequencies (e.g., Hale, 2001; Levy, 2008).

Additionally, the accounts that make reference to language experience in online sentence processing have different claims on how children arrive at their knowledge. For the early abstraction account, it is assumed that children start creating broader generalizations between morphosyntax and thematic roles from the earliest stages of acquisition, (Messenger & Fisher, 2018; Özge et al., 2019), instead of starting with lexically-based or item-based categories before proceeding to abstractions (Ambridge & Lieven, 2011; Tomasello, 2003). According to this approach, acquiring abstract knowledge is not a prolonged and gradual process regardless of whether that knowledge is innate or is rapidly acquired via learning across verbs and constructions. However, a current shortcoming of the approach is that it is formally underspecified. Notably, the content of children's early abstract knowledge is unclear (with assumptions varying widely, see Messenger & Fisher, 2018), as well as how the input changes and interacts with these abstract representations. This underspecification makes it difficult to derive predictions beyond "children are adult-like", which they are frequently not. Focusing on developmental studies of processing is the key to illuminating and explaining these differences, alongside careful analysis of children's input (e.g., Ovans, Huang, & Feldman, 2020) and how their knowledge states vary (Kidd, Donnelly, & Christiansen, 2018).

One possible way to explain the difference in processing between 5-year-olds and 7- and 9-year-olds in Study 3 that is consistent with the frequency-sensitive early abstraction approach is to suggest that children's processing speed gets faster with development (Dick et al., 2004;

Kail & Salthouse, 1994). On this explanation, all children operate with the same grammatical systems, but differ in their ability to rapidly execute parsing routines (Messenger & Fisher, 2018). One possible way to test this suggestion would be to slow the input signal to see if 5-year--old children could predict the upcoming noun in less demanding contexts. However, we point out here that in the experiment the first noun marker + adjective region was, on average, 900 ms long, which is already much longer than it would likely be in naturalistic speech. Furthermore, explaining the developmental effect independent of changes in the grammatical system is inconsistent with research that shows changes in language processing speed in both children and adults reflects existing knowledge and experience (e.g., Donnelly & Kidd, 2020; Fernald, Perfors, & Marchman, 2006; Wells, Christiansen, Race, Acheson, & MacDonald, 2009).

Because of this, we lean towards connectionist accounts that provide a more detailed model of acquisition of grammatical representations and their online use. Connectionist models use error-based learning approaches (Chang et al., 2006; Dell & Chang, 2014; Fitz & Chang, 2019), which entails learning through listening. While listening, the model predicts the next constituent of an utterance by using the distribution; and more importantly, it exploits the differences between the predicted output and the target output (i.e., error), and uses these to update the connection weights which were responsible for the original prediction. This means that the model continues to update across each new experience and creates frequency-based expectations. Thus, upon hearing the frequent patient voice-marked verb, participants have a strong expectation that the *ng* marker would be followed by the agent, and the *ang* marker would be followed by the patient, with the strength of those expectations varying with age. These processes speed up in 7- and 9-year-olds to the point where predictive looks can be made in the patient voice condition before the noun is heard. In adults, these processes become so abstract that the first noun referent no longer needs to be viewed during processing, presumably because the voice and noun markers in concert with an event representation from the preview are sufficient to understand the meaning. In contrast, upon hearing the unexpected agent voice-marked verb, participants had to update their initial expectation for the more frequent patient voice and its associated noun marking configuration, resulting in a slowdown which seemed to cascade to the rest of the sentence regions (Levy, 2008).

Thus, the connectionist approach gives us both a mechanism that acquires the Tagalog voice alternation and derives predictions about how that knowledge will be used in comprehension. Two additional features of the approach deserve mention. Firstly, the Chang et al. (2006) model is sufficiently flexible to acquire language-specific grammatical generalizations that are refined with experience, in addition to predicting language-internal variation in parsing. Thus, it simulates cross-linguistic differences in sentence production across related and unrelated languages (e.g., Chang, 2009; Chang et al., 2015), and predicts variation in parsing choices within languages for functionally similar structures (Yang et al., 2020). Secondly, the flexibility of the learning mechanism means that syntactic representations are dynamic and constantly refined through experience. On this approach, traditional questions regarding the presence or absence of abstract knowledge make way for a more nuanced account of acquisition and processing as an evolving dynamic system.

Conceptualizing children's online processing as experience-driven over more-or-less abstract knowledge has the potential to reconcile some apparent inconsistencies in the past literature. The previous finding that Turkish-speaking children can use case is consistent with the proposal that children operate with abstract categories and form-meaning mappings (Özge et al., 2019). In contrast, the result from German studies showing that children fail to use case to identify thematic roles is inconsistent with an early abstraction account of language acquisition (and by implication, processing, Dittmar et al., 2008; Kröger et al., 2017; Schipke et al., 2012). However, the apparent tensions disappear when we interpret the online processing data within the typological properties and usage frequencies of the language, which each exert an influence on the course of acquisition. In particular, the high frequency of agent-initial sentences in German, coupled with the fact that case-marked determiners carry a high functional load (marking gender and number) and have significant ambiguity throughout the paradigm (including ambiguity in nominative and accusative case in feminine and neuter gender), does not lead to a strong expectation that an accusative-marked first argument is the patient (Dittmar et al., 2008). In contrast, the higher prevalence of sentences with an accusative-marked first noun (patient-initial) in Turkish than in German, and its much more transparent case system (Aksu-Koç & Slobin, 1985), results in both an easier to acquire system of mappings and a lower expectation that early appearing nouns will be marked with the nominative case.

This explanation does not rule out the possibility that children can and eventually do operate with abstract knowledge early in development, but that both the acquisition and use of that knowledge online is necessarily language-specific and experience-dependent. Indeed, even in German, children have been shown to be capable of producing and understanding patient-first sentences at quite young ages under the right circumstances (e.g., Brandt, Kidd, Lieven, & Tomasello, 2009; Sauermann, Höhle, Chen, & Järvikivi, 2011). What it does point to, however, is the need for data from a wider range of typologically-diverse languages, and studies that test multiple age groups to fully appreciate the developmental trajectory of online parsing.

Our results also bear upon a recent proposal in the literature concerning how parsing interacts with the early- versus late-arriving cues to drive acquisition. Pozzan and Trueswell (2015) hypothesized that languages with early-arriving cues in a sentence present a processing and learning advantage because early-arriving cues can guide linguistic analyses, whereas late-arriving cues can only revise initial analyses. Since children have difficulties revising their early parsing commitments (Choi & Trueswell, 2010; Trueswell et al., 1999), the approach predicts an early acquisition of Tagalog voice morphology and noun marking, which are utterance-initial. More importantly, the noun markers commonly occur before the first noun, so the voice-marker in the verb and the noun marker can be used to *guide* thematic role assignment, instead of just merely *revising* it. Our results are partially consistent with this theory, although they fail to account for our online results. Thus, while we found evidence that the children understand the markers, their use of the markers to guide structure building online differed according to voice, as it did in adults. This suggests that a distinction between early- and late-arriving cues cannot be divorced from the distributional properties of those cues in the input.

We did not observe any evidence that children followed a word order strategy in our data, despite the fact that previous studies on Tagalog have provided evidence that, in the agent voice, 5-year-old children interpret the first argument as the agent, regardless of the morphosyntactic markers (Garcia et al., 2019; Garcia, Roeser, & Höhle, 2020). Since there is good evidence in the input for a first noun as agent strategy, such a result would not be unexpected. We suspect that important features of our design reduced any tendency young children had to use a word order strategy. Notably, our use of simple and unambiguous transitive scenes that always mapped onto the audio stimuli enabled participants to map event roles to sentential arguments without the need to choose between potential agents, as in previous studies.

We end with a final comment on the pressing need for data that widens the scope of psycholinguistic theory. On best estimates, we have acquisition data from only 1–2% of the world's 7000 or so languages (Lieven & Stoll, 2010), and adult language processing data from a depressingly lower number than that (Anand, Chung, & Wagers, 2011; Jaeger & Norcliffe, 2009). In the face of rampant language endangerment (Evans, 2010), we are rapidly losing opportunities to study languages that will allow us to build a more comprehensive account of the human language faculty. We put our own data in this category. Verb-initial languages are rare among the world's languages (VAP – 6.9%, VPA – 2.1%, Dryer, 2013), and languages with symmetrical voice systems even more so. However, it was these precise typological features that enabled us to tease apart the predictions of alternative theoretical proposals in the field without the need to use rare or discourse-marked structures (e.g., relative clauses).

6. Conclusion

In the current paper, we investigated Tagalog-speaking adults' and children's online processing of basic transitive sentences. Tagalog is a symmetrical voice language that allows free ordering of nouns marked unambiguously for core thematic roles, yet there are clear distributional differences across both voice and argument order, which enabled us to test the competing predictions of prominent accounts of children's online parsing: the experience-based versus early abstraction accounts. Across three studies, we found evidence to suggest that children's and adults' online parsing was significantly influenced by the distributional properties of the language. The results are most consistent with accounts that heavily consider the influence of input frequency in acquisition.

Acknowledgements

This work was in part supported by the Australian Research Council (CE140100041: CI Kidd). We would like to thank Franklin Chang and the reviewers for their valuable comments on an earlier draft of this manuscript. We are also grateful to Jens Roeser, Sebastian Sauppe, Marisa Casillas, Han Sloetjes, Florian Hintz, Jed Pizarro-Guevara, and Ivan Bondoc for their help in analyzing and interpreting the data; Christina Bergmann, Maarten van der Heuvel, and Kathleen Amora for assisting us in creating the experiments; Angela Legaspi, and Micah Marcial for transcribing and coding the corpus data. Lastly, we are indebted to the families who took part in the corpus study, as well as to the students and faculty of East Bagong Barrio Elementary School and De La Salle Araneta University.

Appendix A. List of experimental sentences in Studies 1 and 3

1. Bite		
Agent-initial	The hardworking mouse was biting the dog last Tuesday.	
	Agent voice	*Kumakagat noong Martes ang masipag na daga ng aso.*
	Patient voice	*Kinakagat noong Martes ng masipag na daga ang aso.*
Patient-initial	The mouse was biting a hardworking dog last Tuesday.	
	Agent voice	*Kumakagat noong Martes ng masipag na aso ang daga.*
	Patient voice	*Kinakagat noong Martes ang masipag na aso ng daga.*
Agent-initial	The hardworking cow was biting the monkey last Tuesday.	
	Agent voice	*Kumakagat noong Martes ang masipag na baka ng unggoy.*
	Patient voice	*Kinakagat noong Martes ng masipag na baka ang unggoy.*
Patient-initial	The cow was biting the hardworking monkey last Tuesday.	
	Agent voice	*Kumakagat noong Martes ng masipag na unggoy ang baka.*
	Patient voice	*Kinakagat noong Martes ang masipag na unggoy ng baka.*
2. Capture		
Agent-initial	The healthy monkey was capturing the cow last Tuesday.	
	Agent voice	*Humuhuli noong Martes ang malusog na unggoy ng baka.*
	Patient voice	*Hinuhuli noong Martes ng malusog na unggoy ang baka.*
Patient-initial	The monkey was capturing the healthy cow last Tuesday.	
	Agent voice	*Humuhuli noong Martes ng malusog na baka ang unggoy.*
	Patient voice	*Hinuhuli noong Martes ang malusog na baka ng unggoy.*
Agent-initial	The healthy mouse was capturing the chicken last Tuesday.	
	Agent voice	*Humuhuli noong Martes ang malusog na daga ng manok.*
	Patient voice	*Hinuhuli noong Martes ng malusog na daga ang manok.*
Patient-initial	The mouse was capturing the healthy chicken last Tuesday.	
	Agent voice	*Humuhuli noong Martes ng malusog na manok ang daga.*
	Patient voice	*Hinuhuli noong Martes ang malusog na manok ng daga.*
3. Lift		
Agent-initial	The brave chicken is lifting the monkey tonight.	
	Agent voice	*Bumubuhat ngayong gabi ang matapang na manok ng unggoy.*
	Patient voice	*Binubuhat ngayong gabi ng matapang na manok ang unggoy.*
Patient-initial	The chicken is lifting the brave monkey tonight.	
	Agent voice	*Bumubuhat ngayong gabi ng matapang na unggoy ang manok.*
	Patient voice	*Binubuhat ngayong gabi ang matapang na unggoy ng manok.*
Agent-initial	The brave turtle is lifting the cat tonight.	
	Agent voice	*Bumubuhat ngayong gabi ang matapang na pagong ng pusa.*
	Patient voice	*Binubuhat ngayong gabi ng matapang na pagong ang pusa.*
Patient-initial	The turtle is lifting the brave cat tonight.	
	Agent voice	*Bumubuhat ngayong gabi ng matapang na pusa ang pagong.*
	Patient voice	*Binubuhat ngayong gabi ang matapang na pusa ng pagong.*
4. Catch		
Agent-initial	The persistent chicken is catching the turtle tonight.	
	Agent voice	*Sumasalo noong Lunes ang matiyagang manok ng pagong.*
	Patient voice	*Sinasalo noong Lunes ng matiyagang manok ang pagong.*
Patient-initial	The chicken is catching the persistent turtle tonight.	
	Agent voice	*Sumasalo noong Lunes ng matiyagang pagong ang manok.*
	Patient voice	*Sinasalo noong Lunes ang matiyagang pagong ng manok.*
Agent-initial	The persistent monkey is catching the cat tonight.	
	Agent voice	*Sumasalo noong Lunes ang matiyagang unggoy ng pusa.*
	Patient voice	*Sinasalo noong Lunes ng matiyagang unggoy ang pusa.*
Patient-initial	The monkey is catching the persistent cat tonight.	

(*continued on next page*)

(*continued*)

	Agent voice	*Sumasalo noong Lunes ng matiyagang pusa ang unggoy.*
	Patient voice	*Sinasalo noong Lunes ang matiyagang pusa ng unggoy.*
5. Chase		
Agent-initial	The caring pig was chasing the dog last Sunday.	
	Agent voice	*Humahabol noong Linggo ang maalagang baboy ng aso.*
	Patient voice	*Hinahabol noong Linggo ng maalagang baboy ang aso.*
Patient-initial	The pig was chasing the caring dog last Sunday.	
	Agent voice	*Humahabol noong Linggo ng maalagang aso ang baboy.*
	Patient voice	*Hinahabol noong Linggo ang maalagang aso ng baboy.*
Agent-initial	The caring chicken was chasing the mouse last Sunday.	
	Agent voice	*Humahabol noong Linggo ang maalagang manok ng daga.*
	Patient voice	*Hinahabol noong Linggo ng maalagang manok ang daga.*
Patient-initial	The chicken was chasing the caring mouse last Sunday.	
	Agent voice	*Humahabol noong Linggo ng maalagang daga ang manok.*
	Patient voice	*Hinahabol noong Linggo ang maalagang daga ng manok.*
6. Cure		
Agent-initial	The sweet dog is chasing the mouse this afternoon.	
	Agent voice	*Gumagamot ngayong hapon ang malambing na aso ng daga.*
	Patient voice	*Ginagamot ngayong hapon ng malambing na aso ang daga.*
Patient-initial	The dog is chasing the sweet mouse this afternoon.	
	Agent voice	*Gumagamot ngayong hapon ng malambing na daga ang aso.*
	Patient voice	*Ginagamot ngayong hapon ang malambing na daga ng aso.*
Agent-initial	The sweet cat is chasing the pig this afternoon.	
	Agent voice	*Gumagamot ngayong hapon ang malambing na pusa ng baboy.*
	Patient voice	*Ginagamot ngayong hapon ng malambing na pusa ang baboy.*
Patient-initial	The cat is chasing the sweet pig this afternoon.	
	Agent voice	*Gumagamot ngayong hapon ng malambing na baboy ang pusa.*
	Patient voice	*Ginagamot ngayong hapon ang malambing na baboy ng pusa.*
7. Drag		
Agent-initial	The cheerful dog was dragging the turtle earlier today.	
	Agent voice	*Kumakaladkad kanina ang masayahing aso ng pagong.*
	Patient voice	*Kinakaladkad kanina ng masayahing aso ang pagong.*
Patient-initial	The dog was dragging the cheerful turtle earlier today.	
	Agent voice	*Kumakaladkad kanina ng masayahing pagong ang aso.*
	Patient voice	*Kinakaladkad kanina ang masayahing pagong ng aso.*
Agent-initial	The cheerful monkey was dragging the chicken earlier today.	
	Agent voice	*Kumakaladkad kanina ang masayahing unggoy ng manok.*
	Patient voice	*Kinakaladkad kanina ng masayahing unggoy ang manok.*
Patient-initial	The monkey was dragging the cheerful chicken earlier today.	
	Agent voice	*Kumakaladkad kanina ng masayahing manok ang pagong.*
	Patient voice	*Kinakaladkad kanina ang masayahing manok ng pagong.*
8. Hit		
Agent-initial	The playful cow is hitting the mouse this Monday.	
	Agent voice	*Pumapalo ngayong Lunes ang mapaglarong baka ng daga.*
	Patient voice	*Pinapalo ngayong Lunes ng mapaglarong baka ang daga.*
Patient-initial	The cow is hitting the playful mouse this Monday.	
	Agent voice	*Pumapalo ngayong Lunes ng mapaglarong daga ang baka.*
	Patient voice	*Pinapalo ngayong Lunes ang mapaglarong daga ng baka.*
Agent-initial	The playful cat is hitting the turtle this Monday.	
	Agent voice	*Pumapalo ngayong Lunes ang mapaglarong pusa ng pagong.*
	Patient voice	*Pinapalo ngayong Lunes ng mapaglarong pusa ang pagong.*
Patient-initial	The cat is hitting the playful turtle this Monday.	
	Agent voice	*Pumapalo ngayong Lunes ng mapaglarong pagong ang pusa.*
	Patient voice	*Pinapalo ngayong Lunes ang mapaglarong pagong ng pusa.*
9. Kick		
Agent-initial	The beautiful pig is kicking the chicken today.	
	Agent voice	*Sumisipa ngayong araw ang magandang baboy ng manok.*
	Patient voice	*Sinisipa ngayong araw ng magandang baboy ang manok.*
Patient-initial	The pig is kicking the beautiful chicken today.	
	Agent voice	*Sumisipa ngayong araw ng magandang manok ang baboy.*
	Patient voice	*Sinisipa ngayong araw ang magandang manok ng baboy.*
Agent-initial	The beautiful monkey is kicking the dog today.	
	Agent voice	*Sumisipa ngayong araw ang magandang unggoy ng aso.*
	Patient voice	*Sinisipa ngayong araw ng magandang unggoy ang aso.*
Patient-initial	The monkey is kicking the beautiful pig today.	
	Agent voice	*Sumisipa ngayong araw ng magandang aso ang unggoy.*
	Patient voice	*Sinisipa ngayong araw ang magandang aso ng unggoy.*
10. Pinch		
Agent-initial	The energetic dog is pinching the monkey this afternoon.	
	Agent voice	*Kumukurot ngayong hapon ang masiglang aso ng unggoy.*
	Patient voice	*Kinukurot ngayong hapon ng masiglang aso ang unggoy.*
Patient-initial	The dog is pinching the energetic monkey this afternoon.	
	Agent voice	*Kumukurot ngayong hapon ng masiglang unggoy ang aso.*
	Patient voice	*Kinukurot ngayong hapon ang masiglang unggoy ng aso.*

(*continued on next page*)

(*continued*)

Agent-initial	The energetic chicken is pinching the pig this afternoon.	
	Agent voice	*Kumukurot ngayong hapon ang masiglang manok ng baboy.*
	Patient voice	*Kinukurot ngayong hapon ng masiglang manok ang baboy.*
Patient-initial	The chicken is pinching the energetic pig this afternoon.	
	Agent voice	*Kumukurot ngayong hapon ng masiglang baboy ang manok.*
	Patient voice	*Kinukurot ngayong hapon ang masiglang baboy ng manok.*
11. Prick		
Agent-initial	The obedient mouse is pricking the cow this Monday.	
	Agent voice	*Tumutusok ngayong Lunes ang masunuring daga ng baka.*
	Patient voice	*Tinutusok ngayong Lunes ng masunuring daga ang baka.*
Patient-initial	The obedient mouse is pricking the cow this Monday.	
	Agent voice	*Tumutusok ngayong Lunes ng masunuring baka ang daga.*
	Patient voice	*Tinutusok ngayong Lunes ang masunuring baka ng daga.*
Agent-initial	The obedient turtle is pricking the dog this Monday.	
	Agent voice	*Tumutusok ngayong Lunes ang masunuring pagong ng aso.*
	Patient voice	*Tinutusok ngayong Lunes ng masunuring pagong ang aso.*
Patient-initial	The turtle is pricking the dog this Monday.	
	Agent voice	*Tumutusok ngayong Lunes ng masunuring aso ang pagong.*
	Patient voice	*Tinutusok ngayong Lunes ang masunuring aso ng pagong.*
12. Pull		
Agent-initial	The loving cow was pulling the pig last Sunday.	
	Agent voice	*Humihila noong Linggo ang mapagmahal na baka ng baboy.*
	Patient voice	*Hinihila noong Linggo ng mapagmahal na baka ang baboy.*
Patient-initial	The cow was pulling the loving pig last Sunday.	
	Agent voice	*Humihila noong Linggo ng mapagmahal na baboy ang baka.*
	Patient voice	*Hinihila noong Linggo ang mapagmahal na baboy ng baka.*
Agent-initial	The loving cat was pulling the monkey last Sunday.	
	Agent voice	*Humihila noong Linggo ang mapagmahal na pusa ng unggoy.*
	Patient voice	*Hinihila noong Linggo ng mapagmahal na pusa ang unggoy.*
Patient-initial	The cat was pulling the loving monkey last Sunday.	
	Agent voice	*Humihila noong Linggo ng mapagmahal na unggoy ang pusa.*
	Patient voice	*Hinihila noong Linggo ang mapagmahal na unggoy ng pusa.*
13. Punch		
Agent-initial	The smart pig is punching the cat today.	
	Agent voice	*Sumusuntok ngayong araw ang matalinong baboy ng pusa.*
	Patient voice	*Sinusuntok ngayong araw ng matalinong baboy ang pusa.*
Patient-initial	The pig is punching the smart cat today.	
	Agent voice	*Sumusuntok ngayong araw ng matalinong pusa ang baboy.*
	Patient voice	*Sinusuntok ngayong araw ang matalinong pusa ng baboy.*
Agent-initial	The smart cow is punching the turtle today.	
	Agent voice	*Sumusuntok ngayong araw ang matalinong baka ng pagong.*
	Patient voice	*Sinusuntok ngayong araw ng matalinong baka ang pagong.*
Patient-initial	The pig is punching the smart cat today.	
	Agent voice	*Sumusuntok ngayong araw ng matalinong pagong ang baka.*
	Patient voice	*Sinusuntok ngayong araw ang matalinong pagong ng baka.*
14. Push		
Agent-initial	The kind mouse was pushing the cat last Monday.	
	Agent voice	*Tumutulak noong Lunes ang mabait na daga ng pusa.*
	Patient voice	*Tinutulak noong Lunes ng mabait na daga ang pusa.*
Patient-initial	The mouse was pushing the kind cat last Monday.	
	Agent voice	*Tumutulak noong Lunes ng mabait na pusa ang daga.*
	Patient voice	*Tinutulak noong Lunes ang mabait na pusa ng daga.*
Agent-initial	The kind turtle was pushing the chicken last Monday.	
	Agent voice	*Tumutulak noong Lunes ang mabait na pagong ng manok.*
	Patient voice	*Tinutulak noong Lunes ng mabait na pagong ang manok.*
Patient-initial	The turtle was pushing the kind chicken last Monday.	
	Agent voice	*Tumutulak noong Lunes ng mabait na manok ang pagong.*
	Patient voice	*Tinutulak noong Lunes ang mabait na manok ng pagong.*
15. Shoot		
Agent-initial	The active pig is shooting the cow tonight.	
	Agent voice	*Bumabaril ngayong gabi ang maliksing baboy ng baka.*
	Patient voice	*Binabaril ngayong gabi ng maliksing baboy ang baka.*
Patient-initial	The pig is shooting the active cow tonight.	
	Agent voice	*Bumabaril ngayong gabi ng maliksing baka ang baboy.*
	Patient voice	*Binabaril ngayong gabi ang maliksing baka ng baboy.*
Agent-initial	The active cat is shooting the mouse tonight.	
	Agent voice	*Bumabaril ngayong gabi ang maliksing pusa ng daga.*
	Patient voice	*Binabaril ngayong gabi ng maliksing pusa ang daga.*
Patient-initial	The cat is shooting the active mouse tonight.	
	Agent voice	*Bumabaril ngayong gabi ng maliksing daga ang pusa.*
	Patient voice	*Binabaril ngayong gabi ang maliksing daga ng pusa.*
16. Tickle		
Agent-initial	The arrogant dog was tickling the pig earlier today.	
	Agent voice	*Kumikiliti kanina ang mayabang na aso ng baboy.*

(*continued on next page*)

(*continued*)

	Patient voice	*Kinikiliti kanina ng mayabang na aso ang baboy.*
Patient-initial	The dog was tickling the arrogant pig earlier today.	
	Agent voice	*Kumikiliti kanina ng mayabang na baboy ang aso.*
	Patient voice	*Kinikiliti kanina ang mayabang na baboy ng aso.*
Agent-initial	The arrogant turtle was tickling the cow earlier today.	
	Agent voice	*Kumikiliti kanina ang mayabang na pagong ng baka.*
	Patient voice	*Kinikiliti kanina ng mayabang na pagong ang baka.*
Patient-initial	The turtle was tickling the arrogant cow earlier today.	
	Agent voice	*Kumikiliti kanina ng mayabang na baka ang pagong.*
	Patient voice	*Kinikiliti kanina ang mayabang na baka ng pagong.*

Appendix B. List of comprehension questions in Study 3

	Agent voice	Patient voice
1. *Kagat* 'Bite'	*Sino ang kumakagat?*	*Sino ang kinakagat?*
	Who is biting?	Who is being bitten?
2. *Huli* 'Capture'	*Sino ang humuhuli?*	*Sino ang hinuhuli?*
	Who is capturing?	Who is being captured?
3. *Karga* 'Carry'	*Sino ang kumakarga?*	*Sino ang kinakarga?*
	Who is carrying?	Who is being carried?
4. *Salo* 'Catch'	*Sino ang sumasalo?*	*Sino ang sinasalo?*
	Who is catching?	Who is being caught?
5. *Habol* 'Chase'	*Sino ang humahabol?*	*Sino ang hinahabol?*
	Who is chasing?	Who is being chased?
6. *Gamot* 'Cure'	*Sino ang gumagamot?*	*Sino ang ginagamot?*
	Who is curing?	Who is being cured?
7. *Kaladkad* 'Drag'	*Sino ang kumaladkad?*	*Sino ang kinakaladkad?*
	Who is dragging?	Who is being dragged?
8. *Palo* 'Hit'	*Sino ang pumapalo?*	*Sino ang pinapalo?*
	Who is hitting?	Who is being hit?
9. *Sipa* 'Kick'	*Sino ang sumisipa?*	*Sino ang sinisipa?*
	Who is kicking?	Who is being kicked?
10. *Kurot* 'Pinch'	*Sino ang kumukurot?*	*Sino ang kinukurot?*
	Who is pinching?	Who is being pinched?
11. *Tusok* 'Prick'	*Sino ang tumutusok?*	*Sino ang tinutusok?*
	Who is pricking?	Who is being pricked?
12. *Hila* 'Pull'	*Sino ang humihila?*	*Sino ang hinihila?*
	Who is pulling?	Who is being pulled?
13. *Suntok* 'Punch'	*Sino ang sumusuntok?*	*Sino ang sinusuntok?*
	Who is punching?	Who is being punched?
14. *Tulak* 'Push'	*Sino ang tumutulak?*	*Sino ang tinutulak?*
	Who is pushing?	Who is being pushed?
15. *Baril* 'Shoot'	*Sino ang bumabaril?*	*Sino ang binabaril?*
	Who is shooting?	Who is being shot?
16. *Kiliti* 'Tickle'	*Sino ang kumikiliti?*	*Sino ang kinikiliti?*
	Who is tickling?	Who is being tickled?

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2021.104859.

References

Abbot-Smith, K., Chang, F., Rowland, C., Ferguson, H., & Pine, J. (2017). Do two and three year old children use an incremental first-NP-as-agent bias to process active transitive and passive sentences?: A permutation analysis. *PLoS One, 12*(10), Article e0186129.

Aksu-Koç, A., & Slobin, D. I. (1985). Acqusition of Turkish. In D. I. Slobin (Ed.), *The crosslinguistic study of language acqusition, Vol. 1: The data* (pp. 839–878). Hillsdale, NJ: Lawrance Erlbaum Associates.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247–264.

Altmann, G. T. M., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*(4), 583–609.

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

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*(2), 239–273.

Ambridge, B., & Lieven, E. V. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press.

Amora, K. K., Garcia, R., & Gagarina, N. (2020). Tagalog adaptation of the multilingual assessment instrument for narratives: History, process and preliminary results. In N. Gagarina, & J. Lindgren (Eds.), *New language versions of MAIN: Multilingual assessment instrument for narratives – revised* (pp. 221–233).

Anand, P., Chung, S., & Wagers, M. (2011). Widening the net: Challenges for gathering linguistic data in the digital age. In *NSF SBE 2020: Future research in the social, behavioral & economic sciences*. Retrieved from https://www.nsf.gov/sbe/sbe_2020/Abstracts.pdf.

Audacity Team. (2015). Audacity(R): Free audio editor and recorder (version 2.0.5) [computer program]. Retrieved from http://audacity.sourceforge.net/.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. Macwhinney, & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 3–73). New York: Cambridge University Press.

Bautista, M. L. S. (2004). Tagalog-English code switching as a mode of discourse. *Asia Pacific Education Review, 5*(2), 226–233.

Billings, L. (2005). Ordering clitics and postverbal R-expressions in Tagalog: A unified analysis? In A. Carnie, H. Harley, & S. A. Dooley (Eds.), *Verb first: On the syntax of verb initial languages* (pp. 303–339). Amsterdam: John Benjamins.

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

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: An investigation of young German and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics, 20*(3).

Casillas, M., Bergelson, E., Warlaumont, A. S., Cristia, A., Soderstrom, M., VanDam, M., & Sloetjes, H. (2017). A new workflow for semi-automatized annotations: Tests with long-form naturalistic recordings of children's language environments. In *Proceedings of the 18th annual conference of the international speech communication association (interspeech 2017)* (pp. 2098–2102).

Chan, A., Yang, W., Chang, F., & Kidd, E. (2018). Four-year-old Cantonese-speaking children's online processing of relative clauses: A permutation analysis. *Journal of Child Language, 45*(1), 174–203.

Chang, F. (2009). Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language, 61*(3), 374–397.

Chang, F., Baumann, M., Pappert, S., & Fitz, H. (2015). Do lemmas speak German? A verb position effect in German structural priming. *Cognitive Science, 39*(5), 1113–1130.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234–272.

Chen, F., Bao, L., Shattuck, R. M., Borja, J. B., & Gultiano, S. (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*, 275–299.

Choi, Y., & Trueswell, J. C. (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology, 106*(1), 41–61.

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

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B, 369*(20120394), 1–9.

Dick, F., Wulfeck, B., Krupa-Kwiatkowski, M., & Bates, E. (2004). The development of complex sentence interpretation in typically developing children compared with children with specific language impairments or early unilateral focal lesions. *Developmental Science, 7*(3), 360–377.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development, 79*, 1152–1167.

Donnelly, S., & Kidd, E. (2020). Individual differences in lexical processing efficiency and vocabulary in toddlers: A longitudinal investigation. *Journal of Experimental Child Psychology, 192*, 104781.

Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer, & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/81.

ELAN (Version 5.9) [Computer software]. (2020). *The language archive*. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://archive.mpi.nl/tla/elan.

Evans, N. (Ed.). (2003). *The non-Pama-Nyungan languages of Northern Australia: Comparative studies of the continent's most linguistically complex region*. Canberra: Pacific Linguistics.

Evans, N. (2010). *Dying words: Endangered languages and what they have to tell us*. Malden, MA: Wiley-Blackwell.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*(1), 98.

Fitz, H., & Chang, F. (2019). Language ERPs reflect learning through prediction error propagation. *Cognitive Psychology, 111*, 15–52.

Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguistic Research, 27*(3), 339–374.

Foley, W. A. (1998, June). Symmetrical voice systems and precategoriality in Philippine languages. In *Paper presented at the 3rd LFG conference*. Brisbane, Queensland, Australia.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91*(4), 1357–1392.

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

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

Garcia, R., & Kidd, E. (2020). The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development, 16*(4), 399–425.

Garcia, R., Roeser, J., & Höhle, B. (2019). Thematic role assignment in the L1 acquisition of Tagalog: Use of word order and morphosyntactic markers. *Language Acquisition, 26*(3), 235–261.

Garcia, R., Roeser, J., & Höhle, B. (2020). Children's online use of word order and morphosyntactic markers in Tagalog thematic role assignment: An eye-tracking study. *Journal of Child Language, 47*(3), 533–555.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684–691.

Good, P. I. (2005). *Permutation, parametric, and bootstrap tests of hypotheses* (3rd ed.). New York: Springer Science & Business Media.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*(4), 274–279.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725.

Hahne, A., Eckstein, K., & Friederici, A. D. (2004). Brain signatures of syntactic and semantic processes during children's language development. *Journal of Cognitive Neuroscience, 16*, 1302–1318.

Hale, J. (2001). *A probabilistic Earley parser as a psycholinguistic model* (Paper presented at the NAACL).

Himmelmann, N. P. (2005a). The Austronesian languages of Asia and Madagascar: Typological characteristics. In K. A. Adelaar, & N. P. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 110–181). London, England: Routledge.

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

Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Linguistic Society of America, 56*, 251–299.

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

Jaeger, T. F., & Norcliffe, E. (2009). The cross-linguistic study of sentence production. *Lang & Ling Compass, 3*(4), 866–887.

Janciauskas, M., & Chang, F. (2018). Input and age-dependent variation in second language learning: A connectionist account. *Cognitive Science, 42*(S2), 519–554.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86*(2–3), 199–225.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133–156.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences, 22*(2), 154–169.

Kidd, E., Stewart, A. J., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language, 38*(1), 222–234.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95*(1), 95–127.

Kröger, J., Münster, K., & Knoeferle, P. (2017). The influence of prosody and case marking on thematic role assignment in ambiguous action scenes: Adults versus children. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual meeting of the cognitive science society* (pp. 2463–2468).

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177.

Lidz, J. (2021). Second-year syntax: Discovering dependencies [keynote address]. In *34th annual CUNY conference on human sentence processing*. https://www.youtube.com/watch?v=HEuYM07mtPM&list=PL1ynN79FU2hUMCvpFSCnC-ku0IiFvLU24&index=3 (March 4-6).

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition, 87*(3), 151–178.

Lieven, E., & Stoll, S. (2010). Language. In M. Bornstein (Ed.), *Handbook of cultural developmental science* (pp. 143–160). New York, London: Psychology Press.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution [corrected]. *Psychological Review, 101*(4), 676–703.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (Ed.). (1999). *The emergence of language*. Lawrence Erlbaum Associates Publishers.

Mahboob, A., & Cruz, P. (2013). English and mother-tongue-based multilingual education: Language attitudes in the Philippines. *Asian Journal of English Language Studies, 1*, 1–19.

Malicsi, J. C. (2013). *Gramar ng Filipino*. Quezon City: Sentro ng Wikang Filipino.

Maris, E. (2012). Statistical testing in electrophysiological studies. *Psychophysiology, 49*(4), 549–565.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190.

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

Mayer, M. (1969). *Frog, where are you?* New York: Dial Books for Young Readers.

Messenger, K., & Fisher, C. (2018). Mistakes weren't made: Three-year-olds' comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition, 178*, 118–132.

Montgomery, J. W., Evans, J. L., Gillam, R. B., Sergeev, A. V., & Finney, M. C. (2015). "Whatdunit?" Developmental changes in children's syntactically based sentence interpretation abilities and sensitivity to word order. *Applied PsychoLinguistics, 37*(6), 1281–1309.

Nordlinger, R. (2011). Transitivity in Murrinh-Patha. *Studies in Language, 35*(3), 702–734.

Ovans, Z., Huang, Y. T., & Feldman, N. H. (2020). The (un)surprising kindergarten path. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 350–356). New York: Curran Associates, Inc.

Özge, D., Kuntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition, 183*, 152–180.

Philippine Statistics Authority. (2003). National Capital Region: Close to 10 million persons (results from the 2000 census of population and housing, NSO). Retrieved from http://www.psa.gov.ph/population-and-housing/node/428.
Philippine Statistics Authority. (2016). Highlights on household population, number of households, and average household size of the Philippines (2015 census of population). Retrieved from https://psa.gov.ph/population-and-housing/node/69728.
Phillips, C. (2013). Parser-grammar relations: We don't understand everything twice. In M. Sanz, I. Laka, & M. Tanenhaus (Eds.), *Language down the garden path: The cognitive basis for linguistic structure* (pp. 294–315). Oxford University Press.
Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism, 5*(4), 409–453.
Phillips, C., & Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. In C. Chesi (Ed.)*, vol. 6*. *Studies in linguistics* (pp. 11–47).
Pozzan, L., & Trueswell, J. C. (2015). Revise and resubmit: How real-time parsing limitations influence grammar acquisition. *Cognitive Psychology, 80*, 73–108.
R Core Team. (2016). *R: A language and environment for statistical computing (Version 3.6.2) [computer program]*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.
Riesberg, S. (2014). *Symmetrical voice and linking in Western Austronesian languages (Pacific linguistics 646)*. Boston, MA: De Gruyter Mouton.
Sauermann, A., Höhle, B., Chen, A., & Järvikivi, J. (2011). Intonational marking of focus in different word orders in German children. In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, & B. Tomaszewicz (Eds.), *Proceedings of the 28th west coast conference on formal linguistics* (pp. 313–322). Somerville, MA: Cascadilla Proceedings Project.
Sauppe, S. (2016). Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences. *Frontiers in Psychology, 7*, 1–13.
Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin, R. D., Jr., & Levinson, S. C. (2013). Dependencies first: Eye tracking evidence from sentence production in Tagalog. In *Paper presented at the the 35th annual meeting of the cognitive science society (CogSci 2013), Austin*.
Schachter, P., & Otanes, F. T. (1972). *Tagalog reference grammar*. Berkeley, CA: University of California Press.
Schipke, C. S., Friederici, A. D., & Oberecker, R. (2011). Brain responses to case-marking violations in German preschool children. *Neuroreport, 22*(16), 850–854.
Schipke, C. S., Knoll, L. J., Friederici, A. D., & Oberecker, R. (2012). Preschool children's interpretation of object-initial sentences: Neural correlates of their behavioral performance. *Developmental Science, 15*, 762–774.
Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147.
Segalowitz, N. S., & Galang, R. G. (1978). Agent-patient word-order preference in the acquisition of Tagalog. *Journal of Child Language, 5*, 47–64.
Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23*(4), 569–588.
SensoMotoric Instruments GmbH. (2017). *BeGaze 3.7: Manual*.
Simons, G. F., & Fennig, C. D. (2018). *Ethnologue: Languages of the world* (21st ed.). Dallas, TX: SIL International. Available from http://www.ethnologue.com.
Skeide, M. A., Brauer, J., & Friederici, A. D. (2016). Brain functional and structural predictors of language performance. *Cerebral Cortex, 26*(5), 2127–2139.
Snedeker, J. (2013). Children's sentence processing. In V. Gompel (Ed.), *Sentence processing* (pp. 189–220). New York: Psychology Press.
Strotseva-Feinschmidt, A., Schipke, C. S., Gunter, T. C., Brauer, J., & Friederici, A. D. (2019). Young children's sentence comprehension: Neural correlates of syntax-semantic competition. *Brain and Cognition, 134*, 110–121.
Tanaka, N. (2016). *An asymmetry in the acquisition of Tagalog relative clauses* (unpublished doctoral dissertation). University of Hawaii at Manoa.
Thompson, R. M. (2003). *Filipino English and Taglish: Language switching from multiple perspectives*. Amsterdam: John Benjamins.
Thothathiri, M., Asaro, C. T., Hsu, N. S., & Novick, J. M. (2018). Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension. *Cognition, 178*, 162–177.
Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language, 58*(2), 188–213.
Tomasello, M. (2003). *Constructing a language: A usage-based account of language acquisition*. Cambridge, MA.
Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89–134.
Tsoukala, C., Broersma, M., van den Bosch, A., & Frank, S. L. (2021). Simulating code-switching using a neural network model of bilingual sentence production. *Computational Brain and Behavior, 4*, 87–100.
Tupas, R., & Lorente, B. P. (2014). A "new" politics of language in the Philippines: Bilingual education and the new challenge of the mother tongues. In P. Sercombe, & R. Tupas (Eds.), *Language, education and nation-building* (pp. 165–180). London: Palgrave Macmillan.
Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology, 58*, 250–271.
Yang, W., Chan, A., Chang, F., & Kidd, E. (2020). Four-year-old mandarin-speaking children's online comprehension of relative clauses. *Cognition, 196*, 104103.
Zhou, P., & Ma, W. (2018). Children's use of morphological cues in real-time event representation. *Journal of Psycholinguistic Research, 47*, 241–260.
Zuckerman, S., Pinto, M., Koutamanis, E., & van Spijk, Y. (2016). A new method for testing language comprehension reveals better performance on passive and principle B constructions. In J. Scott, & D. Waughtal (Eds.), *BUCLD 40: Proceedings of the 40th annual Boston University conference on language development* (pp. 443–456). Somerville: Cascadilla Press.