PROBABILISTIC AND PREDICTIVE PARSING
IN TAGALOG VOICE ALTERNATIONS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

December 2020

Ivan Paul Mungcal Bondoc

Dissertation Committee:
Amy J. Schafer, Chairperson
Kamil Deen
Bradley McDonnell
William O’Grady
Teresita D. Amore
Acknowledgments

I will forever be in gratitude for the endless support and encouragement from many people who have contributed in a significant way in ensuring the successful completion of my dissertation research, and the smooth navigation of my graduate school life. I cannot claim the trophy alone, for it is with several different people that I reached the finish line.

No words can ever express how thankful I am to my dissertation supervisor, Amy Schafer – thank you for honing me to be an independent scholar, for challenging my ideas, and for being an excellent example of a methodical and an ethical researcher. Thank you for steering the direction of my dissertation and my psycholinguistic research agenda, while offering immeasurable ounces of understanding and support in trying times. I am indebted to your patience and wisdom in helping me navigate the space of cognitive-linguistic models and abstract thoughts.

I am also thankful to everyone on my dissertation committee, for their willingness to take part in my committee and helpful comments. Special thanks go to William O’Grady and Kamil Deen, for igniting questions and discussions on syntax, language acquisition, and Philippine-type voice, for their guidance on my research projects, for their continued support on my research pursuits, and for always keeping their doors open for advice.

I offer my heartfelt gratitude to the numerous teachers and mentors I’ve met along the way, for whose teachings and life insights expanded my knowledge and world views, served as a ray of light in troubles and dilemmas, and contributed significantly to my personal and professional growth.

While the depths of knowledge on Philippine-type voice system further remain to be better understood, investigating cognitive-linguistic questions with languages with this feature has been an exciting endeavor, and I would not have completed my experiments without the assistance of different organizations and departments: the Center for Philippine Studies and the Filipino Program at the University of Hawai‘i at Mānoa; Department of Speech Pathology, College of
Allied Medical Professions, University of the Philippines Manila; Department of Linguistics, University of the Philippines Diliman; and UH Pinoy group. My endless thanks go to all participants who devoted their time and resources for my doctoral research.

I am indebted to several people from the Departments of Linguistics, Second Language Studies, East Asian Languages and Literature, Educational Psychology, Shinshu Kyokai Dormitory, East-West Center, and the Newman Center. It has been a meaningful journey we’ve shared for the past five years, and they have been awesome friends and colleagues, with whom I’ve benefitted from the healthy exchange of ideas, perspectives and life stories; for pulling me away from my mental abyss; for lending a hand in times of need; and for being company during exhilarating and tough times. I’ve also treasured the healthy academic discussions and fun times with all of you.

Time and distance have forged my bonds with wonderful friends and colleagues from the University of the Philippines Manila and Diliman, and from Philippine Cultural College, to whom I owe my strength and passion in research. Thank you for being supportive and for offering life insights and words of encouragement when I need it the most. Innumerable conversations with these people always hit home and fueled my passion to keep going. (Many of these friends and colleagues also helped me a great deal with my research, either by participating in my experiments, by agreeing to partake in my endless pilot studies, or by assisting in participant recruitment.)

I am privileged to have been in Hawaii for graduate school. As a settler, the beauty of the islands and its vibrant community have been a source of mental and spiritual enlightenment in endless ways I could imagine. The aloha spirit continually reminds me of our pursuit of universal goodness.

To my family, thank you for your unwavering love and guidance as I trudge through the academic path. We grieved a lot this year, but the memories of those who departed helped us to get through, and they will always be there to remind us how to move forward.
Abstract

Probabilistic and Predictive Parsing in Tagalog Voice Alternations

People tend to predict upcoming elements in language as speech unfolds. Prediction, or the generation of expectations about upcoming input, is argued to be the unifying principle of the human information processing system. Despite the lack of consensus among scholars on its definition and mechanisms, the majority of the extensive work on prediction has focused on generating expectations for a specific lexical item and its lower-level features, giving less attention to the ability of the parser to generate gradient expectations beyond the upcoming word, such as the unfolding linear word order of a sentence.

This dissertation extends our knowledge on prediction by investigating how an understudied type of linguistic information, called voice morphology, is used to develop probabilistic expectations about a sentence’s particular word order. Five experiments examine whether comprehenders use verb and voice morphology early in the sentence to anticipate its likely arguments and predict the likely order of the phrases that follow. I test the hypothesis that Tagalog comprehenders use voice morphology to develop probabilistic syntactic expectations to predict specific sentential word order patterns.

Experiments 1 and 2 investigated the role of agentivity and pivothood in Tagalog word order preferences and on the anticipation of the verb’s likely arguments. A sentence continuation task (Experiment 1) demonstrated the role of two probabilistic and equally-weighted constraints, namely agent-first and pivot-second, in shaping word order preferences across voices, and highlighted the strong link between voice and word order. However, despite this link, comprehenders were not observed to anticipate the agent and the pivot as the likely arguments of the verb in a visual world eyetracking study (Experiment 2); instead, gaze patterns illustrated preferential looks to animates and undifferentiated looks across voices. I speculated these results were a product of comprehenders’ engagement in a wide range of predictions across voices.
Three reading experiments examined whether voice morphology is used by comprehenders to predict the specific linear word order pattern of the sentence, following a gradient cline of patterns. In a self-paced reading study (Experiment 3), comprehenders showed limited predictive effects of voice, as it was only at a later sentence region that they showed the hypothesized gradience of patterns. These limited effects were subsequently verified in two experiments with induced cognitive load. When faced with time pressure, predicted gradience diminished, and comprehenders only generated coarse-grained predictions of the unfolding sentence in a rapid serial visual presentation task (Experiment 4). Correspondingly, limited effects of generating gradient expectations of linear word order were demonstrated in a self-paced reading task with induced memory load (Experiment 5), replicating results from Experiment 3.

In contrast to the major claim that prediction is immediate and ubiquitous (e.g., DeLong et al., 2014; Kuperberg & Jaeger, 2016), these findings reveal the limits of prediction and the variability of prediction in sentence processing. I argue that these results are best captured by a dynamic account that describes the trajectory of probabilistic activation for the unfolding linguistic structures in the course of processing as shaped by gradient prediction and integration.

Keywords: prediction, integration, Philippine-type voice, word order, gradience, expectation, anticipation
Table of Contents

Acknowledgments........................................................................................................ ii
Abstract.......................................................................................................................... iv
Table of Contents.......................................................................................................... vi
List of Tables................................................................................................................ x
List of Figures............................................................................................................... xii
Abbreviations.............................................................................................................. xiii

Chapter 1. Introduction..................................................................................................... 1
1.1. Predictive Processing............................................................................................... 3
   1.1.1. Evidence and Information Sources for Predictive Processing.................... 3
   1.1.2. Prediction at Different Linguistic Levels.................................................... 6
   1.1.3. Computational Formalisms for Prediction............................................... 7
   1.1.4. Issues on Prediction.................................................................................. 10
      1.1.4.1. What is predicted, and how does predictive processing happen?.......... 10
      1.1.4.2. What evidence counts as prediction?............................................. 14
      1.1.4.3. Delineating prediction from other related concepts......................... 15
      1.1.4.4. The pervasiveness of prediction..................................................... 16
      1.1.4.5. Automaticity of predictions.......................................................... 19
   1.1.5. Summary...................................................................................................... 20
1.2. The Tagalog Voice System..................................................................................... 20
   1.2.1. Word Order............................................................................................... 23
      1.2.1.1. Distributional patterns.................................................................... 23
      1.2.1.2. Explanations of word order preferences based on generative accounts........................................................................................................ 25
1.3. Psycholinguistic Work on Tagalog Word Order.................................................. 27
   1.3.1. On Preference and Variation Using Offline Measures................................. 27
   1.3.2. On Online Processing.............................................................................. 29
1.4. Composition of the Dissertation.......................................................................... 31
Chapter 2. Constraint Influences on Linear Word Order in Production
and Anticipatory Gaze Patterns in Comprehension

2.1. Experiment 1: Sentence Continuation Task

2.1.1. Introduction and Hypotheses

2.1.1.1. Pivothood or Syntactic Prominence

2.1.1.2. Agent Prominence

2.1.1.3. Animacy

2.1.2. Hypotheses

2.1.3. Method

2.1.3.1. Participants

2.1.3.2. Materials

2.1.3.3. Procedure

2.1.3.4. Analysis

2.1.4. Results

2.1.4.1. Correct Responses

2.1.4.2. Incorrect Responses

2.1.5. Discussion

2.2. Experiment 2: Visual World Eyetracking

2.2.1. Introduction and Hypotheses

2.2.2. Method

2.2.2.1. Participants

2.2.2.2. Materials

2.2.2.2.1. Critical Items

2.2.2.2.2. Filler Items

2.2.2.3. Procedure

2.2.2.4. Analyses

2.2.2.4.1. Critical Items

2.2.2.4.2. Filler Items

2.2.3. Results

2.2.3.1. Comprehension Accuracy
2.2.3.2. Eye Gaze Data ................................................................. 59
2.2.3.3. Post-hoc Analysis: Critical items versus Fillers ................ 62
2.2.4. Discussion ........................................................................ 66
2.3. General Discussion .............................................................. 71
2.4. Conclusion .......................................................................... 73

Chapter 3. Voice Effects on the Prediction of Word Order Patterns ... 75

3.1. Experiment 3: Self-Paced Reading ......................................... 75
  3.1.1. Introduction and Hypotheses ......................................... 75
  3.1.2. Method ................................................................. 77
    3.1.2.1. Participants ......................................................... 79
    3.1.2.2. Materials: Development of Stimuli Items ................. 79
    3.1.2.3. Materials: Experimental Items ................................. 80
    3.1.2.4. Procedure .......................................................... 80
    3.1.2.5. Analyses .............................................................. 81
  3.1.3. Results ......................................................................... 81
    3.1.3.1. Comprehension Accuracy ................................. 82
    3.1.3.2. Reaction Times .................................................. 83
  3.1.4. Discussion ................................................................... 86

3.2. Experiment 4: Rapid Serial Visual Presentation .................... 89
  3.2.1. Introduction and Hypotheses ......................................... 89
  3.2.2. Method ................................................................. 91
    3.2.2.1. Participants ......................................................... 92
    3.2.2.2. Materials ............................................................ 92
    3.2.2.3. Procedure .......................................................... 92
    3.2.2.4. Analyses .............................................................. 92
  3.2.3. Results ......................................................................... 93
  3.2.4. Discussion ................................................................... 94

3.3. Experiment 5: Self-Paced Reading with Memory Load .......... 98
  3.3.1. Introduction and Hypotheses ......................................... 98
  3.3.2. Method ................................................................. 100
List of Tables

Table 1. Evaluating the permutation of voice and word order patterns against Kroeger’s principles (1993)…………………………………………………………………………………24

Table 2. Results of the mixed effects logistic regression models that tested the likelihood of the agents being the first argument, and the pivots being the second argument in Experiment 1………………………………………………………………………………….45

Table 3. A subset analysis of the responses with either agents or patients in the first and second positions in the agent voice and patient voice conditions, using mixed effects logistic regression in Experiment 1……………………………………………………………………..46

Table 4. Counts of elements appearing in the first and second argument positions in Experiment 1…………………………………………………………………………………………..50

Table 5. Percent use of the intended agents/pivots in participants’ sentence completions in Experiment 1………………………………………………………………………………………54

Table 6. Percentage of correct responses in the post-stimulus task across conditions in Experiment 2…………………………………………………………………………………………59

Table 7. Results of mixed effects linear regression models that tested for a preference to look at the pivot-associated AOI during the critical verb + adverb region, for each AOI in Experiment 2……………………………………………………………………………………61

Table 8. Results of the linear regression models that tested the proportion of looks to animates versus inanimates in Experiment 2……………………………………………………63

Table 9. Results of the linear regression models that examined the looks to patients/benefactives as they were mentioned in NP2/NP3 in PV and BV sentences in Experiment 2…………………………………………………………………………64

Table 10. Results of the linear regression models that examined the looks to Animate1/Inanimate4 as they are mentioned in NP2/NP3 in a subset of unaccusative sentences in Experiment 2…………………………………………………65

Table 11. A list of the stimulus items for each condition in Experiment 3……………………78

Table 12. Results of the mixed effects logistic regression models for comprehension accuracy in Experiment 3……………………………………………………………...83
Table 13. Results of the mixed effects linear regression models that tested the reaction time differences between conditions, per sentence region in Experiment 3

Table 14. Results of the mixed effects logistic regression models that examined comprehension accuracy between the two RSVP conditions (Experiment 4) and the post-stimulus task in self-paced reading (Experiment 3)

Table 15. Results of the mixed effects logistic regression models for memory load accuracy in Experiment 5

Table 16. Results of the mixed effects logistic regression models for comprehension accuracy in Experiment 5

Table 17. Results of the mixed effects linear regression models that tested the reaction time differences between conditions, per sentence region in Experiments 3 and 5

Table 18. Results of the mixed effects linear regression models that tested the reaction time differences between conditions per sentence region in Experiment 5
List of Figures

Figure 1. A sample item in the online sentence completion task in Experiment 1 ..........41

Figure 2. The number of responses with correctly marked arguments for each voice type in Experiment 1 .........................................................43

Figure 3. Frequency of the different arguments in the first (left panel) and second position (right panel) in correct responses in Experiment 1 .................................44

Figure 4. Frequency of the different arguments in the first (left panel) and second position (right panel) in incorrect responses in Experiment 1 ..................................47

Figure 5. Visual scenes for example filler sentences 5 (top-left), 6 (top-right), 7 (bottom-left), and 8 (bottom-right) in Experiment 2 ..............................55

Figure 6. Proportion of looks to AOIs depicting the agent, benefactive, instrument, and patient arguments, by voice condition in Experiment 2 .......................60

Figure 7. Proportion of looks to animate/inanimate entities in the filler items, by sentence type in Experiment 2 ..................................................62

Figure 8. Subset analysis involving the looks to mentioned/unmentioned elements in selected sentence regions for the critical items (left) and the filler items (right) in Experiment 2 ..................................................64

Figure 9. Comprehension accuracy rates in the post-stimulus task in Experiment 3 ....82

Figure 10. Reading times (log-transformed) in the four conditions across the sentence regions in Experiment 3 ..................................................83

Figure 11. Accuracy rates in the two voice and word order conditions across the two RSVP tasks (Experiment 4) and the self-paced reading (SPR) task (Experiment 3) ..................................................93

Figure 12. Hypothesized reaction times in Regions 3 and 5 in Experiment 5 ...............100

Figure 13. Accuracy rates in the memory load task across conditions in Experiment 5 ....103

Figure 14. Comprehension accuracy rates in the different voice and word order conditions between Experiments 3 and 5 .......................................104

Figure 15. Reading times (log-transformed) in the four conditions in Experiments 3 and 5 across the sentence regions .................................................106
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3PL</td>
<td>third person plural</td>
</tr>
<tr>
<td>3SG</td>
<td>third person singular</td>
</tr>
<tr>
<td>AG</td>
<td>agent argument</td>
</tr>
<tr>
<td>AV</td>
<td>agent voice</td>
</tr>
<tr>
<td>BEN</td>
<td>benefactive argument</td>
</tr>
<tr>
<td>BV</td>
<td>benefactive voice</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive argument</td>
</tr>
<tr>
<td>IMP</td>
<td>imperfective</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INS</td>
<td>instrument argument</td>
</tr>
<tr>
<td>INT</td>
<td>intransitive</td>
</tr>
<tr>
<td>IV</td>
<td>instrument voice</td>
</tr>
<tr>
<td>LNK</td>
<td>linker</td>
</tr>
<tr>
<td>LOC</td>
<td>locative argument</td>
</tr>
<tr>
<td>LV</td>
<td>locative voice</td>
</tr>
<tr>
<td>MULTAG</td>
<td>multiple agents</td>
</tr>
<tr>
<td>NPVT</td>
<td>non-pivot</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>PAT</td>
<td>patient argument</td>
</tr>
<tr>
<td>PCV</td>
<td>verb of perception</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PRF</td>
<td>perfective</td>
</tr>
<tr>
<td>PV</td>
<td>patient voice</td>
</tr>
<tr>
<td>PVT</td>
<td>pivot</td>
</tr>
<tr>
<td>REL</td>
<td>relativizer</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

The idea of prediction, or the generation of expectations about upcoming input, takes a dominant role in the current literature on sentence comprehension. Initially rejected in the 1970s and 80s as a central principle of sentence processing, it has made a resurgence in the literature over the past two decades, with different scholars having variable definitions, mechanisms, and claims for its role in the information processing system (e.g., Altmann & Mirković, 2009; DeLong et al., 2014; Huettig, 2015; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018). Most models of prediction generally assume it has a pervasive role across various levels of representation, with most research centered on the prediction of a lexical item and its sub-lexical and semantic features (e.g., Altmann & Kamide, 1999; DeLong et al., 2005; Federmeier & Kutas, 1999b; Kamide, Altmann, et al., 2003; Knoeferle et al., 2005; Laszlo & Federmeier, 2009; Luke & Christianson, 2016; Staub et al., 2015). A smaller number of studies have investigated syntactic prediction, or the generation of expectations for the likely unfolding phrasal structure (e.g., Lau et al., 2006; Staub, 2010; Staub & Clifton, 2006; Wicha et al., 2003). From a contemporary computational perspective, prediction involves the generation of expectations at a word-by-word level as a function of the noisiness of the accrued sentential context or of the developing memory representation (Futrell et al., 2020; Futrell & Levy, 2017; Gibson et al., 2013).

The current knowledge on syntactic prediction has been greatly shaped by research evidence from languages with fixed word orders, which tend to restrict the hypothesis space about the upcoming syntactic category or nominal argument. It remains poorly understood whether the parser can formulate expectations beyond the upcoming word, especially when grammatical information is available early on and is probabilistically informative of the likely word order of a sentence.

This dissertation aims to sharpen our understanding of prediction and expectation-based syntactic comprehension by investigating how an understudied type of linguistic information, called voice morphology, is used to develop probabilistic expectations for a sentence’s particular word order. Tagalog, a verb-initial language, provides an opportunity to examine this question. Tagalog verbs contain voice morphology that potentially has implications on the likely types and
order of arguments that one may anticipate. Despite the lack of consensus on the canonical word order patterns in Tagalog, there are claims in the literature that sentential patterns do not occur in equal frequency across voices, and that word order may be strongly correlated with voice morphology (Billings, 2005; Kroeger, 1993). Each type of voice morphology suggests the occurrence of sentential word order patterns in various probabilities. Hence, it can be assumed that the sentence parser may utilize these statistical frequencies relative to voice morphology to develop specific expectations for the particular word order. An examination of Tagalog can test the ability of the parser to generate gradient syntactic expectations based on the preferences for certain voice patterns. These voice preferences would be informative of the degree of processing cost associated with generating expectations for the sentence.

Tagalog also establishes a good testing ground for prediction as its flexible word order creates a test for the generation of syntactic expectations in a more expansive probabilistic space, for instance, the material beyond the next word. Compared to verb-medial or verb-final languages, Tagalog is less restrictive in terms of the ordering of its nominal arguments. Doing an investigation on Tagalog permits us to examine the parser’s ability to use grammatical information to formulate successful predictions at a distance (such as the probable second argument after a verb), or to demonstrate evidence for the unsuccessful cases of immediate prediction, which would call for further refinement of prediction models.

This dissertation examines the major hypothesis that prediction occurs cross-linguistically from all types of linguistic information by examining the understudied role of Tagalog voice alternations in sentence processing. Specifically, I test the hypothesis that Tagalog comprehenders use voice morphology to develop probabilistic syntactic expectations to predict specific sentential word order patterns. With the following experiments, I show that comprehenders demonstrate weak predictions of the gradient linear word order patterns, suggesting that predictive processing should be reframed under a dynamic framework that describes the outcomes of pre-activation under variable strengths as the sentence temporally unfolds.

This chapter proceeds as follows. First, I review the literature that offers evidence for predictive processing in sentence comprehension, and discuss the current issues related to prediction. I then provide an overview of the Tagalog voice system and controversies on Tagalog word order preferences and variation. A discussion follows on the experimental and psycholinguistic work
that delves into word order preferences and the processing of Tagalog word order. This background chapter concludes by providing an overview of the chapters that follow in this dissertation.

### 1.1. Predictive Processing

#### 1.1.1. Evidence and Information Sources for Predictive Processing

Early studies considered the possible value of prediction in processing the upcoming words in a sentence (Miller & Isard, 1963; Tulving & Gold, 1963). However, several researchers were hesitant with the idea that the processing system merely develops predictions in sentence comprehension. They argued that prediction may be too computationally costly and the hypothesis space would be too expansive for the system to predict upcoming structure out of all possible alternatives (Forster, 1981; Jackendoff, 2002). Sentence processing models were then dominated by the concept of integration, the retroactive process of linking constituents into units to efficiently disambiguate and extract meaning from language form. Integration-based models focused then on the resolution of syntactic ambiguities, as well as the roles of memory and interference in sentence processing (Frazier, 1979; Frazier & Clifton, 1989; Frazier & Fodor, 1978; Gibson, 2000; MacDonald, Just, & Carpenter, 1992; MacDonald, Pearlmutter, & Seidenberg, 1994). Yet, the current state of the field has once again experienced a paradigm shift, with a growing interest in prediction as a critical component of models of sentence comprehension.

There is an abundance of research utilizing the visual world eyetracking paradigm to investigate prediction in comprehension, with the major finding that different types of information are used by comprehenders to generate their predictions. Several studies have examined how comprehenders use semantic and syntactic information to anticipate the unfolding signal. Altmann and Kamide’s (1999) seminal work revived the interest on the role of prediction in comprehension. Their study found that people begin to anticipate the likely grammatical object of a verb before it unfolds in the linguistic signal. An example visual scene in this experiment is a picture of a boy sitting in a room with a truck, cake, train, and ball. Thus, when people hear the sentence “The boy will eat…,” they access the semantic properties of the verb *eat* and then restrict their fixations to only the edible entity in the visual context and identify it as a potential grammatical object of the verb, such as the entity *cake* in their study (Altmann & Kamide, 1999). Subsequent research has found how a comprehender’s knowledge of a verb’s indirect object (Kamide, Altmann, &
Haywood, 2003), verb subcategorization information (Arai & Keller, 2013), and case marking (Kamide, Altmann, et al., 2003; Kamide, Scheepers, et al., 2003) are used to make predictions on upcoming input.

Prosody has been reported to influence predictions as well. Using visual world eyetracking, Weber, Grice, and Crocker (2006) investigated whether German listeners utilize prosodic information to anticipate the post-nominal argument in structurally ambiguous SV- and OV-sentences. German SVO sentences tend to have a nuclear pitch accent on the verb, in contrast to the OVS sentence where the nuclear pitch accent occurs on the preverbal nominal. The results of their experiment showed German speakers constraining their looks to either the post-verbal -O or -S argument, contingent with the prosodic pattern presented in the sentence. Other studies have shown similar findings on the use of prosodic information to predict upcoming input (Nakamura, Arai, & Mazuka, 2012; Snedeker & Trueswell, 2003).

Linguistic information can also be coupled with contextual information to drive predictive looks on unfolding input. A study by Kaiser and Trueswell (2004) examined whether the presence of context, together with case marking information, can establish the discourse status (given/new) of an element, and drive anticipatory looks on the upcoming discourse-relevant element. This question was investigated with a flexible word order language, Finnish. In an SVO structure, neither the subject nor the object is informative of the discourse status. In contrast, the elements in a dispreferred OVS structure disambiguates discourse status: the object only refers to a given element, while the subject refers to new information. Finnish comprehenders were observed to anticipate the discourse-new element prior to the completion of the unfolding post-nominal argument in the OVS compared to the SVO pattern. In a similar vein, non-linguistic factors, such as visual context and real world knowledge were also observed to influence the prediction of upcoming input. Knoeferle, Crocker, Scheepers, and Pickering's (2005) research showed how disambiguating visual context can drive anticipatory looks to the upcoming element in structurally ambiguous SV- and OV- structures in German. The influence of visual context and real-world knowledge on predicting the unfolding input were observed in other work (Altmann & Kamide, 2007; Chambers et al., 2004; Ferreira, Foucart, & Engelhardt, 2013; Knoeferle & Crocker, 2006, 2007; Spivey et al., 2002).
Neuropsychological methods have also brought to light predictive evidence in comprehension. Predictive evidence from classic research that measured semantic anomalies using event-related potentials (ERP) found varying amplitudes of the N400 component relative to the degree of fit of a word to the semantic context, as in the likely continuation of different verb types in the sentence “The pizza was too hot to... eat/drink/cry.” The amplitude of the N400 was inversely related to the probability of people using a word as a likely continuation of a sentence frame, called cloze probabilities (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980, 1984). These cloze probabilities are often computed by determining the number of individuals that had a particular completion for that cloze within that sample (Taylor, 1953).

The finding of a reduced N400 component has been more recently associated with prediction. DeLong, Urbach, and Kutas (2005) probed how English speakers predicted the upcoming nominal argument by examining the article (a/an) alternation in English. They predicted that speakers would expect an article congruent with the likely continuation of the sentence fragment (e.g., The day was breezy so the boy went outside to fly... a kite/an airplane). Looking at the ERP responses at the time speakers encounter the article directly examines prediction rather than integration, they argued, given that there is no semantic difference between the articles that can pose integration difficulty. The study outcomes showed a reduced N400 component when an article is congruent with the expected noun than when it is not. In addition, the amplitude in the N400 component was found to be correlated with the likelihood of the upcoming noun, suggesting that people do probabilistic pre-activation of elements prior to encountering the noun. Related work has reported the difference in ERP measures on articles, determiners, or classifiers, depending on its congruence with the gender or category of the upcoming noun (Kwon et al., 2017; Van Berkum et al., 2005; Wicha et al., 2003).

Other neuropsychological studies have described the strength of prediction based on the pre-activation of the semantic features of the upcoming nominal. Federmeier & Kutas (1999a) examined how the degrees of semantic fit of the sentence continuation elicited gradations on the N400 component. Given a sentence like “They wanted to make the hotel look more like a tropical resort. So, along the driveway, they planted rows of...” participants either encountered a continuation from a highly expected word (e.g., palms), an unexpected word from the expected semantic category (e.g., pines), or an unexpected word from an unexpected semantic category
(e.g., tulips). The outcomes revealed not only an ERP difference between the expected and the unexpected words, but also an ERP difference between the unexpected words from the two different semantic categories. Federmeier and Kutas interpreted the difference between the two unexpected words as resulting from the pre-activation of semantic features of the upcoming words, rather than the difficulty to integrate the encountered word from the context. This pre-activation of semantic features of the upcoming words characterizes the anticipation of an upcoming element. Similar findings of semantic pre-activation as a function of cloze probability have been reported in related work (Federmeier et al., 2007; Federmeier & Kutas, 1999a; Staub et al., 2015).

There is compelling evidence for prediction based on reading measures. Classic work by Ehrlich and Rayner (1981) has shown that more predictable words based on the previous context are more likely to be skipped and have reduced fixated time on in reading than less predictable ones. Other studies contribute to the growing support for the independent influence of cloze predictability in early eye reading measures, along with word length (Balota et al., 1985; Brysbaert et al., 1998; Rayner et al., 2011), context-independent word frequency (Rayner et al., 1996), and contextual constraint (Frisson et al., 2017; Staub et al., 2015). Staub’s critical review (2016) finds the predictability effect to be an argument for predictive rather than integrative processing, as these early reading measures are more likely to reflect pre-lexical or early lexical pre-activation of the upcoming word, rather than the later reading measures that are associated with integration difficulties. These pre-activations that happen at a pre-lexical or early lexical stage were found to take effect even before contextual support influences the processing, hence contributing to the role of prediction in processing (Brothers et al., 2015).

1.1.2. Prediction at Different Linguistic Levels

Prediction at various linguistic levels has been described to happen at multiple levels of representation. In the previous section, examples of lexical-semantic prediction have been presented (DeLong et al., 2005; Federmeier et al., 2007; Federmeier & Kutas, 1999b; Kutas & Hillyard, 1984; Staub, 2015), but pre-activation of upcoming syntactic structure, as well as phonological and orthographical form, also exist.

Evidence for the immediate prediction of syntactic structure comes from literature that investigated syntactic pre-activation in a variety of syntactic phenomena. Staub and Clifton (2006)
demonstrated how readers had faster reading times on coordinated NPs or independent clauses when the word *either* was available early in the sentence, such as in “The workers painted *(either)* the house or the barn over the summer” and “*(Either)* our neighbor shoveled the snow or plowed the driveway”, suggesting the use of the previously encountered word (either…) to pre-activate the upcoming conjoining phrase (…or…). Lau et al. (2006) found an early left anterior negativity (ELAN) response in noun phrases where ellipsis would have been expected by a comprehender, as in “*Although Erica kissed Mary’s mother, she did not kiss Dana’s (of the bride).*” Other studies have found predictive effects for parasitic gaps in subject island effects (Phillips, 2006), relativization (Staub, 2010), sluicing (Yoshida et al., 2013), and grammatical gender agreement (Wicha et al., 2003), among others.

The classic study discussed above by DeLong et al. (2005) on the anticipation of the most likely determiner (a/an) illustrates the pre-activation of some of the phonological features of the upcoming nominal, i.e., the onset of the lexical item (vowel/consonant). Kim and Lai (2012) found ERP effects suggesting the pre-activation of orthographical forms in upcoming lexical items. Comprehenders showed a P130 to target pseudowords that orthographically resembled a real word compared to real words (e.g., *She measured the flour so she could bake a ceke/cake*), while pseudowords that were plausible (*tont*) or implausible (*srdt*) manifested a different effect (an N170). Similar effects of pre-activation of the phonological/orthographical forms were also reported in other studies (Laszlo & Federmeier, 2009). However, in some cases, more time was needed to observe successful effects of phonological/orthographical pre-activation (Ito et al., 2016), suggesting that while pre-activation could occur at various levels of representation, the timing of successful pre-activations could differ at various levels, as the probability space for the alternative candidates for the semantic and syntactic representations could be smaller compared to the alternatives for the lexical and phonological features (Kuperberg & Jaeger, 2016; Staub, 2015).

1.1.3. Computational Formalisms for Prediction

Prediction has been computationally formalized under several accounts (Kuperberg & Jaeger, 2016) in terms of the construct of *expectations*. As the system receives a linguistic signal, it develops probabilistic expectations for an upcoming word based on the statistical properties of the grammatical system and the context. A related measure is an element’s *predictability* or
expectancy, which is the probability for an element to occur in a particular sentence frame or context. In more recent work, the reaction times in cloze tasks have been established as a product of a target word’s probability and the degree of item constraint to that target word (Staub et al., 2015).

An equivalent metric for expectations is surprisal, which is computed as the negative log-probability of a word relative to the context where it occurs. Surprisal reflects the degree to which the prediction generated by the system has been met or not met (Hale, 2001; Levy, 2008). Surprisal (Levy, 2008; N. J. Smith & Levy, 2013) tends to be considered by these researchers a typical measure of predictability (DeLong et al., 2014; Huettig, 2015; Huettig & Guerra, 2019; Van Petten & Luka, 2012).

These expectation-based accounts are often used to explain processing difficulty: assuming that context is controlled, expectations that are met induce lower surprisal rates (and thus lower processing difficulty), while dashed expectations incur higher surprisal rates (and thus higher processing cost). These cases of dashed expectations can be observed in garden-path or syntactically ambiguous sentences that have been widely attested in the literature, such as the classic example *The horse raced past the barn fell* (Frazier, 1979; Frazier & Fodor, 1978; MacDonald et al., 1994).

In addition, as the system continues to incrementally receive the input, the conditional probabilities for the predicted upcoming input are adjusted accordingly (Hale, 2001; Levy, 2008). Within the context of Bayesian formalisms, predictive comprehension proceeds as a continuous cycle of belief updating. These generated expectations function as the prior distribution, which will be updated upon the accrual of new information (Doya et al., 2007).

Psycholinguistic evidence has shown that expectation-based accounts are useful in explaining processing difficulty that occurs in less frequent or less canonical patterns in the language, as in the case of subject/object relative clause alternations. As a case in point, Staub (2010) has shown higher surprisal rates at the nominal inside object relative clauses (e.g., senator in (2)) compared to its counterpart in subject relative clauses. Elevated surprisal rates were observed at the object relative clause verb compared to the subject relative clause verb.

(1) The reporter [that attacked the senator] admitted the error. (SRC)

(2) The reporter [that the senator attacked] admitted the error. (ORC)
Other studies have found how people rapidly develop syntactic expectations for the upcoming structure, such as the case of either...or... patterns (Staub & Clifton, 2006), ellipsis (Lau et al., 2006), or sluicing (Yoshida et al., 2013), as discussed in the previous section.

Subsequent studies on expectation-based accounts have often been evaluated against memory-based accounts. In contrast to expectation-based accounts, memory-based accounts describe the processing cost of integrating a word from the sentence context based on the distance and number of unfilled dependencies as each word is obtained (e.g., Gibson, 2000); some describe the processing cost as a function of interference between elements bearing similar features (e.g., Lewis et al., 2006; Wagers et al., 2009). A study on Russian relative clauses, a language that permits flexibility in word order, has shown the need for both expectation- and memory-based components (Levy et al., 2013). Further studies in the asymmetry of relative clauses in verb-final languages have also served as a good test for evaluating these two accounts. In some of these verb-final languages, expectation-based accounts extend support for an SRC advantage due to its statistical frequency, while memory-based accounts tend to identify an ORC advantage due to the shorter dependency distance between the head of the relative clause and the verb. Research on Korean (Kwon et al., 2010, 2006) and Japanese (Miyamoto & Nakamura, 2003; Ueno & Garnsey, 2008) both favor a subject relative clause advantage, expressing support for expectation-based accounts. In contrast, Mandarin, a non-verb-final language, initially showed conflicting results between a subject/object relative clause advantage (Hsiao & Gibson, 2003; Vasishth, Chen, Li, & Guo, 2013). Later evidence has shown value for expectation-based accounts for Mandarin (Hsiao & MacDonald, 2016; Jäger, Chen, Li, Lin, & Vasishth, 2015).

More recently, expectation-based accounts have incorporated memory-based components in noisy/lossy-context models in their computational formalisms, as these describe the system’s limits to parse the noisy input from the context and generate a 100% accurate memory representation of the sentence (Futrell et al., 2020; Futrell & Levy, 2017; Gibson, Piantadosi, et al., 2013; Hahn et al., 2020). The extension of the term expectations and surprisal will be limited in this dissertation to refer to Levy's (2008) seminal account on expectation-based surprisal.
1.1.4. Issues in Prediction

Predictive processing has not been devoid of controversies, despite its abundant evidence in the literature. Some scholars cast doubt on the role of prediction in comprehension. This section revisits each critical issue.

1.1.4.1. What is predicted, and how does predictive processing happen?

Researchers differ in defining the scope of prediction and the linguistic representations that are predicted by the cognitive system. In general, most researchers broadly refer to prediction as the pre-activation of any linguistic input prior to receiving the linguistic signal (Huettig, 2015; Huettig & Guerra, 2019). Under their account, the scope of prediction is extensive, such that it encompasses pre-activation at any level of linguistic representation.

Other scholars adhere to specific representations and/or mechanisms that are involved in predictive processing. Most of these scholars are neuropsychological researchers who instantiate the inference of an upcoming specific lexical item as prediction, while the immediate, fast-acting activation of some semantic content, such as semantic, lexical, phonological, or even perceptual features, is only considered to be (neural) pre-activation, expectation or anticipation (DeLong et al., 2014; Van Petten & Luka, 2012). Following this definition, prediction is seen as an all-or-none, deterministic process, while pre-activation or anticipation is a gradient process.

To other scholars, prediction is linked with memory retrieval mechanisms, such as the pre-activation of the stored memory representations prior to the upcoming input (Chow, Momma, et al., 2016; Chow et al., 2018). In particular, their “bag-of-words” approach to prediction illustrates how a measure of prediction such as the N400 only indexes comprehenders’ superficial pre-activation of the upcoming lexical item, as influenced by the lexical-semantic relationships of the words in the previously accrued input.

Altmann and Mirković’s (2009) account embodies how prediction is utilized as a means to realize the mapping between language form and an event representation which corresponds to some real-world event or knowledge. Prediction in their account is defined in two senses using connectionist accounts. The first definition of prediction operationalizes the process as an innate mechanism of the system, with the goal of activating an output at time $t$ relative to the unfolding bottom-up input that would occur at time $t + 1$. The second definition of prediction defines it as
the *internal state*, an ability of the system to activate units at the hidden layers from the input acquired at time $t - 1$. The term *anticipation* is interchangeably used with this latter definition of prediction. Prediction under Altmann and Mirković’s account operates at varying grain sizes, depending on which linguistic level is involved. Both linguistic and non-linguistic input, as well as the internal state of the system, constitutes the context that serves as the driving force for predictive processing.

Several accounts of prediction also emphasize the generation of higher-level representations in influencing pre-activation at lower levels of representation. To Kim and colleagues (2016), a higher-level representation, such as structured event knowledge, plays a critical role for prediction. Kuperberg and Jaeger's (2016) *active generative* account extensively discusses the role of higher-level representation in prediction.¹ Their definition includes the rational use of higher-level information to actively generate a probabilistic inference at the same higher-level representation, prior to receiving the unfolding bottom-up linguistic signal. The goal of the rational agent to infer the higher-level representation with a greater degree of certainty facilitates pre-activation at lower levels of representation. Kuperberg and Jaeger expand on these perspectives to incorporate in their account of prediction the relevance of Bayesian priors, *surprisal* as an index of anticipatory processing, cycles of belief updating, and the rational or goal-oriented tendency of agents to actively generate predictions in real-time (Yan et al., 2017).

Ferreira & Chantavarin (2018) believe in the critical role of higher-level discourse representations in the comprehension process. They link integration as a complementary process with prediction in comprehension. They define integration as the linking of information that the system has accrued. Instead of prediction, they prefer the use of the term *preparedness* in describing the extent by which we become receptive to the unfolding signal, the result of which is the pre-activation at various linguistic levels. They argue that the developing discourse representations are used in preparation for the unfolding signal, and are enriched via integration processes.

Similarly, some scholars foreground the production system as a key component of the mechanistic process for prediction. Pickering and Garrod (2007) believe that prediction is a crucial piece of the puzzle in interweaving the production and comprehension systems. In their account,

¹The term *generative* should not be confused with Chomsky’s generative syntax framework.
comprehenders are in a dynamic process of predicting and imitating covertly across different levels of representation, with the production system serving as an emulator. An algorithm decides in an incremental manner whether top-down prediction or bottom-up input influences processing. A strong prediction and a noisy input results in the use of top-down information, while a weak prediction and a clear input results in the use of bottom-up information. Lupyan and Clark’s (2015) prediction model for the general perceptual system has a resemblance to Pickering and Garrod’s (2007) model, in that the system engages in either top-down (predictive) and bottom-up (input) flow of information. The goal of processing is to minimize prediction error; the cognitive system’s decision to put more weight on either the top-down or bottom-up process depends on the nature of the task.

Pickering and Garrod’s later account of prediction (Pickering & Garrod, 2013) incorporates the notion of dialogue coordination, and compares production and comprehension to action and action perception. The goal for the processing system is to ensure a coordinated dialogue between interlocutors, and prediction could either proceed by simulation or association. When the system makes predictions by simulation, the interlocutors’ dialogue must be in harmony and the production system is used to generate forward models of action and action perception. In contrast, predictions that proceed by association occur in situations when the interlocutors do not necessarily align in dialogue. That is, when people share the same mental discourse, the listener predicts by covertly imitating and generating the likely output of the speaker. However, if their mental discourse does not overlap, the listener predicts the unfolding signal via priming mechanisms.

A more recent account by Pickering and Gambi (2018) employs prediction-by-production and prediction-by-association as a mechanistic explanation for prediction. Prediction-by-production is an optional mechanism by the system whereby an individual (a comprehender) is assumed to be covertly imitating the accrued signal, generating a likely higher-level representation based on the speaker’s intentions, and using this higher-level representation to predict the unfolding signal. As a resource-intensive process, prediction-by-production is described to result in more accurate predictions. In contrast, prediction-by-association is an obligatory, automatic process. It augments the prediction-by-production mechanism through the process of cascading activation. Both prediction-by-production and prediction-by-association are described to work hand-in-hand in comprehension.
Dell and Chang's (2013) P-chain model linked prediction with the production system. In their model, language processing incorporates prediction at multiple levels of representation. The anticipation of lower-level representations using higher-level information is likened to the generation of a linguistic message in the production system. The process of generating predictions can be error-prone, and error signals become an opportunity for the system to be primed or to be modified to decrease instances of error in the future. They argue that the latter is a type of implicit learning by which language is acquired.

MacDonald (2013) utilizes the language production system as being relevant for the predictive processes that occur in comprehension. In her Production-Distribution-Comprehension (PDC) account, certain processing tendencies in production, such as the preference for structures with greater accessibility, frequency, and prominence, as well as structures with reduced interference shape the distribution of the language form. The comprehension system generates predictive expectations based upon the statistical properties of the language form, which has been influenced by the production system.

There is also a mention in the literature of the use of multiple mechanisms for predictive processing. Huettig (2015) argued that no one mechanism would be adequate to explain the complexity of predictive processing, and as such, the mind is engaged with the use of multiple interacting mechanisms, including production, association, combinatorial systems, and event simulations. These mechanisms are interwoven with the situational context and mediating factors such as age and literacy. The use of these multiple mechanisms has been shown in some studies (Hintz et al., 2016, 2017; Kukona et al., 2011).

To summarize, there has been variation in terms of how scholars define prediction, and what mechanisms are employed to generate predictions. Some scholars limit the scope of prediction to the lexical level, while some extend prediction at various levels of representation. Predictive processing has been framed as either an instantiation of pre-activation reflecting memory access, or as a generation of a higher-level representation that results in cascading activation at lower levels of representation, or as a result of a variety of mechanisms.
1.1.4.2. What evidence counts as prediction?

Researchers have differed with regard to what types of research evidence count as prediction. Many scholars count research findings to be evidence of prediction if the index of processing is measured before the critical word is encountered by the comprehender (Huettig, 2015; Huettig & Guerra, 2019; Pickering & Gambi, 2018). Kutas and colleagues (2011) expanded this further by classifying models as being either non-predictive, weakly predictive, and strongly predictive, depending on the index of processing, as well as the extent to which the model has the capacity to assume prevalent anticipatory processing. According to them, the information-theoretic models proposed by Hale (2001) and Levy (2008), whilst employing probabilistic inferences on the upcoming input, are non-predictive, as the measure of real-time processing (i.e., surprisal) occurs at a point when the critical word is seen. “Pruning” models engage in the probabilistic selection of a limited set of candidate sentence parses (Jurafsky, 1996; Narayanan & Jurafsky, 2002). These models were described as weakly predictive since the index of processing (i.e., attention shifts as reflective in reading times) only happens at certain sentence points and in limited syntactic cases. “Optimal preparation” models such as those by Smith and Levy (2013) were described as strongly predictive since their predictability measures that happen at arbitrary points in connected texts reflect the cognitive state’s constant engagement in the anticipatory processing of the upcoming signal. Their incremental anticipatory model aimed to increase processing efficiency, while having trade-offs in processing resources.

However, Kuperberg and Jaeger’s (2016) account employs an entirely different perspective. Earlier psycholinguistic models are considered predictive, in the sense that these models generate a syntactic expectation on the unfolding input. Hence, garden-path models (Frazier, 1979; Frazier & Clifton, 1989; Frazier & Rayner, 1982) are interpreted as deterministic models of prediction, while lexicalist models of syntactic ambiguity resolution (MacDonald et al., 1994; Seidenberg & MacDonald, 1999; Trueswell et al., 1993) are gradient models of prediction. Under Kuperberg and Jaeger’s account, all information-theoretic models (Hale, 2001; Levy, 2008), “pruning” models (Jurafsky, 1996; Narayanan & Jurafsky, 2002) and “optimal preparation” models (N. J. Smith & Levy, 2013) all instantiate evidence of gradient prediction, since their assumption of the cognitive system’s constant cycles of updating probability distributions result in new Bayesian priors even before the upcoming signal is encountered.
1.1.4.3. Delineating prediction from other related concepts

Predictive processes are often strongly linked with other related processing concepts. Many scholars have distinguished prediction from priming or passive association mechanisms. While prediction and priming are often described via cascading or spreading activation mechanisms, the major distinction lies at the timing of the activation relative to the input, and the levels of representations involved in the process. Priming or passive association is lingering activation when the input has been received. Priming typically refers to spreading activation within a single level of representation. Studies that have often described lexical priming effects often describe the spreading of activation within a network of related lexical items, once the lexical item has been received (Bentin et al., 1985; Meyer & Schvaneveldt, 1971). Syntactic priming effects can be construed as activation for a certain type of construction that is recently encountered in the input (Branigan et al., 1995; Pickering & Branigan, 1998, 1999; Potter & Lombardi, 1998). In contrast, prediction is anticipatory activation: it describes the activation for a specific linguistic representation, prior to actually receiving that input. Typically, multiple levels of representation are involved with the process of generating pre-activations for a specific linguistic representation, such as the higher levels influencing the lower levels.

In contrast, integration is described as the incorporation of newly acquired information with previously acquired input. A belief-update is a consequence of integrative processes (DeLong et al., 2014; Ferreira & Chantavarin, 2018; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018).

Scholars also differentiate between pre-activation and pre-updating. To some, pre-activation reflects a commitment to one or more likely candidates in the upcoming input, while pre-updating refers to the use of that pre-activation to update the higher-level representation in working memory. Hence, pre-activation precedes the process of pre-updating. This definition was used to delineate pre-activation from priming, whereby the former engages in commitments, while the latter does not (Lau et al., 2013; Ness & Meltzer-Asscher, 2018). Other models slightly differ with regard to the order of these processes, such that pre-updating refers to the use of multiple information sources in the prior distribution to update the higher-level representation, while pre-activation refers to the cascading activation at the lower levels of representation as an influence of the higher-
level representation. In active generative frameworks such as these, pre-updating occurs first prior to pre-activation (Kuperberg & Jaeger, 2016).

Research work has elucidated on the indices of pre-activation in language comprehension, vis-à-vis passive mechanisms and integration. Kuperberg (2007) advocated for two separate but highly interacting streams in language comprehension, occurring in parallel: one stream that involves processing based on lexical-semantic associations, and another stream that involves the combinatorial fit of these elements, as well as the lexical thematic relationships of a verb and its arguments. The first stream is reported to be indexed by the N400, while the second stream is indexed by the P600. Furthermore, Brothers et al. (2015) demonstrated the rapid onset of successful lexical pre-activation in varying degrees of contextual support with reduced amplitudes of N250, suggesting that these pre-activations most likely influence the early stages of lexical processing. Staub's (2015) review of lexical predictability on eye movements in reading suggests that the earliest reading measures, such as word skipping, are correlated with lexical predictability in reading. These early reading measures that are more plausibly associated with early lexical processing imply that they are most likely effects of prediction, and not integration. Similarly, Nieuwland and colleagues' (2018) replication of the DeLong et al. study (2005) have showed two different effects in the N400, to be discussed in further detail in the subsequent section. Thus, both prediction and integration processes are expected to work hand-in-hand in real-time comprehension, with prediction being carried out prior to integration (see subsequent section regarding a separate discussion on replicability issues found in the same study).

1.1.4.4. The pervasiveness of prediction

The role of prediction in the language processing system has been in the spotlight in recent years. Several researchers have described the value of prediction for learning (Conway et al., 2010; Dell & Chang, 2013; Misyak et al., 2010) and in harmonizing the mental discourse of the interlocutors in a dialogue (Ferreira & Chantavarin, 2018; Pickering & Garrod, 2007, 2013). This perceived vital role of prediction has led researchers to believe that prediction is unifying the operations of the human information processing system (Bar, 2009; Clark, 2013; Friston, 2010; Lupyan & Clark, 2015). Some authors have even described human brains as being “prediction machines” (Clark, 2013).
However, some scholars cast doubt on the ubiquity of prediction in language processing. One argument against prediction in processing is the use of “prediction-encouraging” contexts and experimental set-ups in research (Huettig, 2015; Huettig & Mani, 2016). Huettig (2015) argues against the idea that the processing system routinely engages in prediction, given that most of the visual world eyetracking research offering predictive evidence tended to have long preview times and slower-than-normal speech rates. He adds that most of the experimental stimuli utilized rich, semantic context which facilitated predictive effects. In an attempt to probe on these issues, Huettig and Guerra (2019) implemented a study involving visual world eyetracking to examine whether certain set-ups of experimental conditions facilitated prediction. They manipulated the experimental conditions that were favorable for prediction to occur: (1) speech rates of linguistic stimuli; (2) length of preview time; and (3) type of instructions given to the participant. Their study tested whether Dutch speakers use determiners as a cue to anticipate the targeted gender-agreeing nominal in their experimental conditions. Their experimental results showed a lack of predictive effects in normal speech conditions with short preview times. They reported only small anticipatory effects even in conditions where participants were explicitly instructed to predict. They argued that these findings challenge the idea of prediction as a unifying principle of the human information processing system.

A second argument against prediction is the questionable interpretation of the behavioral measures used as evidence for prediction. In particular, Huettig & Guerra (2019) questioned the interpretation of the N400 component, which has often been attributed to prediction. They argued that the N400 component may also be a reflective measure of integration, especially when its measurement was made during, and not prior to, the target word. He added that the studies can only count toward predictive evidence if there was a clever manipulation of the stimuli, such as measuring N400 effects in words that have equal semantic content (e.g., a determiner), before the target nominal is encountered (e.g., DeLong et al., 2005).

To shed light on this issue, Nieuwland and colleagues (2018) re-investigated the N400 effect by conducting a large-scale replication of DeLong and colleagues' study (2005). Their study developed multiple measures to disentangle the conflating effects of predictability, plausibility, and semantic similarity in cloze values as possible correlates for predictive evidence. They tested the hypothesis that an association between predictability and N400 amplitude measures would
reflect prediction, while an association between plausibility and N400 measures would be evidence of integration. The outcomes of their study revealed that both predictability and plausibility had an N400 effect, in that the former contributed to the N400 amplitude rise, while the latter had an effect once the former had set in and had a downward flank. They interpreted the results as showing that the N400 would not be an explanation of prediction alone, and would be best described as a composite of multiple but related cognitive processes (but see Mantegna et al., 2019, for a related study with a different finding). Other studies failed to replicate the results of the classic DeLong et al. (2005) study (Ito et al., 2017; Nieuwland, Politzer-Ahles, et al., 2018). Similarly, Van Petten & Luka (2012) opens up the need to inspect other ERP components that may be relevant to explaining the prediction and/or integration phenomenon during language comprehension.

Additionally, there is evidence from reading studies that readers may not be engaged in ubiquitous prediction. Luke and Christianson's (2016) large-scale cloze task on connected texts of varying lengths demonstrated that only 5% of content words and 19% of function words were highly predictable. This finding suggested that highly predictable words are rare in connected texts; hence, the prediction of upcoming lexical items is unlikely during normal reading. This was complemented by an eyetracking while reading study, which showed that despite the low percentage of highly predictable words in connected texts, readers do not demonstrate processing costs based on reading measures (e.g., longer reading times) as a result of mispredictions. This major finding conflicts with the idea that successful predictions are immediately implemented by the processing system. Because it is extremely rare to predict a specific lexical item, their findings highlighted that gradient prediction or anticipation would be more feasible compared to all-or-none lexical prediction.

In addition, there has been work in the literature that described the variability of comprehenders with regard to predictive processing, which is suggestive of the non-ubiquity of prediction as a fundamental process of the language processing system. Age is one of these influential factors. Older individuals were observed to demonstrate less predictive processing compared to younger adults, although processing performance was observed to be modulated by individual differences (Bornkessel-Schlesewsky et al., 2015; Federmeier et al., 2010; Wlotko et al., 2012). Literacy was observed to influence the capacity to predict. Skilled and more literate readers demonstrated better use of contextual information to make predictions on upcoming input, compared to less literate
readers (Mishra et al., 2012; Ng et al., 2017; for a review, see Huettig & Mishra, 2014), with reading skill as a strong correlate of predictive processing (Mani & Huettig, 2014). Processing speed and working memory were found to mediate predictive processing (Huettig & Janse, 2016). Predictive processing was also reported to be limited among individuals with reading disabilities such as dyslexia (Huettig & Brouwer, 2015), and language disorders such as aphasia (Mack et al., 2013, although see Milburn et al., 2016; Warren et al., 2016) compared to their neurotypical counterparts. Furthermore, majority of the work on psycholinguistic processing focus on a small set of the world’s languages, and tend to recruit participants from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Henrich et al., 2010; Huettig, 2015), thus limiting the generalizability of the study findings. These findings suggest that comprehenders may not necessarily be consistently employing predictive mechanisms in order to comprehend language, in contrast to theories proposed in the literature (e.g., Clark, 2013; DeLong et al., 2014).

1.1.4.5. Automaticity of predictions

Another related issue is the automaticity of predictions in sentence processing. Some scholars describe predictive processes as automatic and largely unconscious (e.g., Bar, 2009; DeLong et al., 2014). Other researchers disagree, suggesting that the malleability of prediction to several cognitive strategies poses a question with its automaticity. If pre-activation mechanisms occur more unconsciously, it should not be sensitive to purposeful processing decisions that require conscious attention and goal-orientedness (Moors & De Houwer, 2006). Previous neuropsychological research has shown support for the sensitivity of predictions to conscious, goal-oriented behaviors or strategies. Semantic priming experiments have shown the degree to which comprehenders adjust their expectations based on the semantic content of the context (Becker, 1980; Neely, 1976). Specifically, Lau et al. (2013) manipulated the proportion of prime-target pairs in a semantic category probe detection task to examine its eventual influence on primes and targets throughout the experiment. Comprehenders showed reductions of the N400 amplitude in conditions with increased semantic relatedness proportions. Similar findings of sensitivity to experimental proportion manipulations have been found as well in sentential cloze contexts. Brothers et al. (2017) manipulated the proportion of item-level sentence frames with a highly predictable target word within an experiment. Their experiment demonstrated how comprehenders
modulate their anticipatory processing (as reflected in reading time measures) depending on the
global statistical properties of the experimental context. Comprehenders made adjustments on their
predictions based on the proportion of highly predictable vs unpredictable target words in
sentential cloze contexts presented in an experiment, as well as when they were instructed to actively predict the upcoming word.

Likewise, syntactic adaptation, or the rational adjustment of comprehenders to the statistical
properties of the accrued input, also serves as evidence to the sensitivity of prediction to cognitive
strategies. Syntactic adaptations have been observed in reading times in self-paced reading
experiments (Fine et al., 2010; Kleinschmidt et al., 2012; Myslin & Levy, 2016), as well as N400 adaptations in ERP experiments (Delaney-Busch et al., 2019).

Kuperberg and Jaeger (2016), however, expressed a word of caution about linking control with the
notion of automaticity, as there are also cases of automatic processes that are sensitive to conscious and goal-oriented behaviors.

1.1.5. Summary

Predictive processing has accumulated evidence from research that spans two decades, where various methods such as visual world eyetracking, neuropsychological methods involving event related potentials, and eyetracking while reading have been utilized. While there is general consensus that predictive processing occurs at multiple linguistic levels and that different types of information can be used to generate predictions, there has been disagreement among scholars with regard to the definitions, nature, and mechanisms of predictive processing. This dissertation aims to shed light on some of these issues, to be tackled more in detail in Section 1.4. The next section introduces the reader to the Tagalog voice system.

1.2. The Tagalog Voice System

As mentioned earlier, most of what has been established about prediction in comprehension has been based on literature from Indo-European languages, such as English, or East Asian languages, such as Mandarin, Japanese, and Korean (e.g., Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003; Knoeferle & Crocker, 2006, 2007; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Spivey, Tanenhaus, Eberhard, & Sedivy,
These languages either have a verb-medial or verb-final word order, with the canonical position of the subject occurring in the first argument position.

The verb initiality and verb system of Tagalog distinguishes it from these more frequently studied languages. Tagalog is a major language in the Philippines spoken by 23 million people worldwide (Eberhard et al., 2020). It employs the Philippine-type voice system, characterized by the presence of voice morphology on the verb that endows a major syntactic privilege to one of the arguments, i.e. the pivot, also referred to in the literature as the focused element, privileged syntactic argument, trigger, topic, primary argument, nominative, absolutive, or (albeit more controversially) subject. In these languages, various thematic roles (e.g., agent, patient, locative, benefactive, instrument, etc.) can be marked as the pivot (Foley, 1998; Himmelmann, 2005; McDonnell, 2016; Riesberg, 2014; Schachter & Otanes, 1972).

This voice system is illustrated in example sentences (3) and (4). The sentences contain the verb stem luto ‘to cook’, which can select various thematic roles as the pivot. Sentence (3) is in the agent voice, where the verb takes an agent voice affix [mag-], selecting the agent nanay ‘mother’ as the syntactically prominent argument or the pivot, marked by the nominal marker ang [ʔaŋ]. The non-prominent (non-pivot) patient spaghetti is marked with ng [naŋ]. Sentence (4) is in the patient voice, where the verb takes the null patient voice affix, and identifies the patient spaghetti as pivot, as indicated by the nominal marker ang. The non-pivot agent nanay is marked with ng. In sentences (3) and (4), the benefactive and instrument receive other nominal marking; however, it is possible to mark one of these elements as the pivot, as is shown in (5) with the benefactive voice, and (6) with the instrument voice.

---

2Note that the nature of this system is still controversial; researchers in the field have used various labels to describe this complex system, such as focus system, voice system, and Philippine-type system, among many others. I use the terms “voice” and “pivot” in this paper as descriptive labels for this system, and not as an analysis of the alignment system of Tagalog.

3I follow historical and theoretical treatments of Tagalog voice affix glossing (Maclachlan 1992; Kroeger 1993; Rackowski 2002; Blust & Chen 2017) that analyze the suffix -in as the patient voice affix, and the infix -in- as the perfective affix. In the perfective aspect of the patient voice, the infix is retained, while the suffix is deleted.
Agent Voice (AV)

(3) **Nag-luto** kani-kanina lang **ang nanay** ng spaghetti para sa anak gamit **ang** kawali.

_Patient Voice (PV)_

(4) **∅-luto** kani-kanina lang ng **nanay ang spaghetti** para sa anak gamit **ang** kawali.

_Benefactive Voice (BV)_

(5) **Ip-luto** kani-kanina lang ng **nanay ang anak** ng spaghetti gamit **ang** kawali.

_Instrumental Voice (IV)_

(6) **Ip-luto** kani-kanina lang ng **nanay ang kawali** ng spaghetti para sa anak.

---

4There are three agent voice affixes in Tagalog: [mag-], [um-], and [maN-] (Himmelmann, 2008; Pittman, 1966; Ramos, 1974; Schachter & Otanes, 1972). Some verbs only take one of these affixes (e.g., mag-laba but *I<um>aba / *man-laba ‘to wash’; t<um>tugtog but *mag-tugtug / *man-tugtug ‘to play (an instrument)’; mang-aso [from maN-aso] but *mag-aso / *umaso ‘to hunt’). Others allow either of these affixes to be used for the agent voice, although with semantic differences between them (e.g., s<um>sulat ‘to write’, mag-sulat ‘to write continuously or intently’; mag-walis ‘to hunt’, mag-walis ‘to sew’, ma-nahi [from maN-tahi] ‘to sew (a number of things), or professionally’).

These three affixes also differ with regard to aspectual paradigms, as summarized in the table below. PR stands for partial reduplication of the verb stem.

<table>
<thead>
<tr>
<th>AV affix</th>
<th>infinitive</th>
<th>past</th>
<th>present</th>
<th>future</th>
</tr>
</thead>
<tbody>
<tr>
<td>[mag-]</td>
<td>[mag-]</td>
<td>[mag-]</td>
<td>[mag-] + PR</td>
<td>[mag-] + PR</td>
</tr>
<tr>
<td>[maN-]</td>
<td>[maN-]</td>
<td>[naN-]</td>
<td>[naN-] + PR</td>
<td>[maN-] + PR</td>
</tr>
<tr>
<td>[um-]</td>
<td>[um-]</td>
<td>[um-]</td>
<td>[um-] + PR</td>
<td>PR</td>
</tr>
</tbody>
</table>

Pittman (1966), Ramos (1974), and Schachter & Otanes (1972) categorized Tagalog verbs depending on the agent voice affixes they take, and they also described the semantic functions these affixes express. For example, Pittman (1966) and Ramos (1974) stated that [mag-] is often used for ‘centrifugal’ transitive (moving an entity away from the agent) and reflexive verbs, while [um-] is used for ‘centripetal’ transitive (moving an entity toward the agent) and nontransient verbs indicating a change of state. Schachter and Otanes (1972) describe [maN-] verbs as those that express activities that are deliberate, intensive, or repetitive toward some entity. Interested readers are directed to these sources for a complete list of these functions.

The critical items for Experiments 1 and 2 utilized verbs that take [mag-] for the agent voice, while the ones in Experiments 3 to 5 take any of these three affixes for the agent voice.

5Pronounced and spelled as ‘niluto’ after a phonological alternation (metathesis). For the purposes of showing the affix, the verb form prior to metathesis is presented in the example.

6The nasal of the instrumental prefix ipang- undergoes place assimilation to the consonant initial of the verb stem. Hence ipang- appears as ipam- before a verb stem with a bilabial initial, ipan- with a dental/alveolar initial, and ipang- with a velar initial. This phonological alternation was reflected in the verb forms in the study.

---

22
'The mother cooked spaghetti for the child with the pan about a while ago.'

While the voice system allows for multiple semantic roles to be chosen as the pivot, these voice alternations differ in terms of their frequency of use. The agent and patient voice patterns are the two most common patterns in the language, with the patient voice occurring more frequently than the agent voice (Cooreman et al., 1984). The benefactive and instrument voice patterns are less frequently used (Himmelmann, 1987).

1.2.1. Word Order

What adds to the intricacy of Tagalog syntax is the flexibility of word order (Schachter & Otanes, 1972). In principle, nominal arguments can be positioned in any order, although not all word order patterns occur with equal frequency. In these sentences, only the noun phrases are switched; nominal marking remains the same relative to voice morphology.

1.2.1.1. Distributional patterns

There is a lack of consensus on the canonical or preferred word order in Tagalog (Himmelmann, 2005; Kroeger, 1993; Schachter, 2015). Kroeger (1993) proposed that despite the word order flexibility in the language, there are three relevant principles that influence the ordering of nominal arguments in Tagalog: (1) agents tend to precede all other arguments; (2) pivots tend to follow all other arguments; and (3) heavier NPs tend to follow lighter NPs.\(^7\) He also added that adverbs and adjuncts tend to occur clause-finally.

The first two of Kroeger’s principles are relevant in the analysis of transitive sentences that involve two arguments, i.e., an agent and a patient. Table 1 summarizes how these principles serve as “driving forces” in Tagalog word order preferences. Table 1 lists the two common voice alternations in Tagalog, crossed with two possible word order patterns (Agent-Patient and Patient-Agent patterns). The first agent voice pattern (AV-Agent-Patient) aligns with the first principle, but not the second one, while the other pattern (AV-Patient-Agent) aligns with the second but not the first; hence, these two patterns could be predicted to be equally preferred. In the patient voice,\(^7\)

---

\(^7\)In Kroeger’s original work, the terms ‘actor’ and ‘R-expression marked nominative’ were used to refer to the ‘agent’ and ‘pivot,’ respectively. For the purposes of consistency throughout this dissertation, the two latter terms will be used.
both principles map perfectly with the strongly preferred PV-Agent-Patient word order; in contrast, neither of these principles align for the dispreferred PV-Patient-Agent pattern. The final column in Table 1 summarizes these tendencies, highlighting that while there is only a single pattern strongly preferred for the patient voice, two patterns would be equally preferred in agent voice. My first dissertation experiment examined the controversy on Tagalog word order preferences and illustrated how these preferences are supported by these predictions.

**Table 1.** Evaluating the permutation of voice and word order patterns against Kroeger’s principles (1993). The pivot element is in boldface for every condition.

<table>
<thead>
<tr>
<th>Voice</th>
<th>Word Order</th>
<th>Word Order Principles</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Principle 1 (on agents)</td>
<td>Principle 2 (on pivots)</td>
</tr>
<tr>
<td>AV</td>
<td>V-Agent-Patient ✓</td>
<td>X</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>V-Patient-Agent X</td>
<td>✓</td>
<td>Yes</td>
</tr>
<tr>
<td>PV</td>
<td>V-Agent-Patient ✓</td>
<td>✓</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>V-Patient-Agent X</td>
<td>X</td>
<td>No</td>
</tr>
</tbody>
</table>

Billings (2005) concurred with Kroeger’s principles by describing parallels of these word order preferences in pronominals. He added that the overwhelming number of patient voice sentences with pronominal arguments occurring in texts demonstrates a general preference for agent pronominals to occur immediately after the verb. Despite the absence of an equivalent set of non-pivot patient pronominals in agent voice, a flexible ordering between an agent and a patient pronominal in agent voice can be similarly observed in certain non-verbal predicates that do not indicate voice.

Other linguists have expressed additional views on the word order issue. Schachter (2015) stated a tendency for agents to precede all other elements, although he did not expound on the reasons for this preference. Himmelmann (2005), on the other hand, mentioned a tendency in Philippine-type symmetrical voice languages for pivots to follow all other arguments (which he refers to as a subject-final tendency in his work, referring to pivots as subjects). He added that non-pivot core elements immediately follow the verbal predicate and form a VP constituent in Western Austronesian languages. Datives tend to occur before the pivot, he states, while other adjuncts

---

8Himmelmann (2005) addresses how the terms *core* and *oblique argument* are defined across scholars. For Tagalog, Himmelmann applies Kroeger’s (1993) distinction of *core* (*ang-* and *ng*-phrases) versus oblique
tend to follow after the pivot element. Kaufman (2009) and Hsieh (2016) take a similar view with Kroeger, which allows for two competing word order preferences in the agent voice.

Richards' (2017) investigation of the interaction between prosody and word order has shown that a similar pattern of pitch rises is observed with verbs immediately followed by thematic subjects compared to thematic objects, irrespective of voice. His work, however, did not discuss the potential implications of the prosodic patterns on word order preferences. In a more recent study, Riesberg, Malcher and Himmelmann (2019) describe word order patterns for several languages with a similar system to Tagalog. They describe two basic word order patterns for Tagalog: a universal agent-first tendency applies to undergoer (non-agent) voice patterns, while a pivot-final tendency applies to the agent voice. As in Himmelman (2005), this predicts patient-agent word order for the agent voice.

1.2.1.2. Explanations on word order preferences based on generative accounts
Several scholars explain the acceptability of word order variation and even the preference to produce certain forms using generative accounts (classifiable under the general Principles and Parameters framework). Guilfoyle, Hung & Travis (1992) expounded on the observable word order variants in Tagalog using Government and Binding Theory. Their two-structural-subject account argued for two structural subject positions, the [Spec, VP] and the [Spec, IP] positions. The verb assumes a position at INFL. The head INFL could either assign case to the [Spec, IP] position via specifier-head agreement, or case-license an element in the [Spec, VP] position through government. The equal preference for two word order patterns in agent voice is attributed to either the movement of the agent in the [Spec, IP] position, which permits an Agent-Patient word order; or, the simple positioning of the agent at the [Spec, VP] position, which produces the Patient-Agent word order. Meanwhile, the single and strongly preferred Agent-Patient word order in the patient voice is caused by a limited option for the patient pivot to be in the [Spec, IP] position. These representations do not account for the less frequently attested patterns in the language, as well as the probabilities of the different word order patterns.

(sa- and other prepositional phrases) in distinguishing these two arguments. This core versus oblique contrast will be sufficient for the purposes of this dissertation.
Rackowski (2002) offers a different explanation of word order variation via post-merger scrambling. With her configurational analysis of Tagalog under the Minimalist framework, she argued for the movement of the pivot to the [Spec, VoiceP] position to allow for the interpretation of specificity. All elements below the Voice head receive a non-specific interpretation. Scrambling follows after the attainment of this basic hierarchical structure, allowing for the flexibility of word order patterns in both voice alternations. These representations, however, do not reflect the frequency differences of the agent voice with the others.

Aldridge (2004) accounts for the word order variation in agent voice via TP-fronting to a [Spec, CP] position. Analyzing Tagalog through an ergative-absolutive lens using the Minimalist framework, she explained a typical base generation of agent voice with an Agent-Patient word order. However, a possible movement of the agent out of its base position triggers TP-fronting, hence leading to the production of a Patient-Agent word order. In patient voice, the agent is projected at a lower [Spec, vP], while the patient moves to a higher [Spec, vP] position for case checking. When these arguments remain in-situ in their case-checked positions, this produces a Patient-Agent order; however, an alternative of having them spelled-out at their base-generated position produces an Agent-Patient word order. While Aldridge’s representations capture the flexibility of word order, they do not explain the relative frequencies of the different word order patterns.

Manueli (2010) takes a similar view as Aldridge (2004) in terms of Tagalog hierarchical structure but argues for an Agent-Patient pattern as the basic word order in Tagalog. Her claim, based on Dryer’s (2007) criteria on the determination of a language’s basic word order, identifies the Agent-Patient pattern being the most basic, owing to its sheer frequency in a variety of voice alternations involving both pronominal and nominal arguments. She further added that a deviation in this word order pattern, specifically the patient voice Patient-Agent word order, causes ambiguities in terms of meaning, and is therefore avoided.

Manueli’s study collected native speaker judgments on the degree of grammaticality of sentences with various word order patterns. The participants in her study judged all agent-first sentences as strongly grammatical, while the non-agent first sentences were judged less grammatical than the agent-first ones, but not as ungrammatical. It is unclear, however, how the grammaticality of these sentences relates to their frequencies in the language.
Chen (2017) proposed a classification of Philippine-type languages based on the word order patterns they exhibit. These include languages that manifest (1) a pivot-final word order, (2) an agent-first word order based on thematic hierarchy, and (3) flexible word order patterns. She classified Tagalog under languages that tended to have an agent-first word order, based on elicited data. With her nominative-accusative analysis of Tagalog under the Minimalist framework, the underlying hierarchical structure of Tagalog involves the external argument (agent) receiving nominative case, and the internal argument (patient) receiving accusative case. The argument that agrees with the voice morphology is promoted to pivot at the [Spec, CP], and this system overrides the underlying case of the pivot. To explain word order variation, Chen elaborated that an overt promotion of an argument to pivot at [Spec, CP] produces the pivot-last word order pattern, while a covert promotion-to-pivot at the LF leads to the Agent-Patient word order. Chen’s representations are similar to Aldridge’s in that they account for structural flexibility; however, they do not account for the relative frequencies of the different word order patterns.

In general, work from these generative accounts have assumed a base-generated Agent-Patient word order, and suggested operations (such as movement) to allow for word order variations. While these accounts provide insight on possible structural representations, and thus capture the range of possible orders, they do not fully account for the descriptive preferences discussed in the literature, that is, the relative frequencies of word orders. The following section discusses the perspectives that psycholinguistic work offers on the observed word order variation preferences in Tagalog.

1.3. Psycholinguistic Work on Tagalog Word Order

1.3.1. On Preference and Variation Using Offline Measures

Very limited psycholinguistic work has explored the production of word order variation preferences in Tagalog.

Hsieh (2016) utilized a read-aloud production task to investigate word order preferences and the prosodic phrase structure of Tagalog transitive declaratives. Sixteen native Tagalog adults were asked to read sentences aloud and rate the naturalness of sentences using a scale of 1 to 7. Participants’ naturalness ratings showed a strong preference for the Agent-Patient word order in
patient voice, but an equal preference between the Agent-Patient and Patient-Agent word order patterns in the agent voice.

Other studies have investigated the linguistic performance of populations such as native Tagalog children and adults with language disorders, along with native Tagalog adults as a baseline. The following reports the performance of native Tagalog adults in this line of research.

A segment of a study by Tanaka (2015; 2016) explored the production of declarative sentences by native Tagalog speakers using a declarative clause elicitation task. Two sets of pictures involving transitive events were presented to the participants: an agent voice-oriented condition, which involved an animate agent and an inanimate patient; and a patient voice-oriented condition, in which both agents and patients were animate. In the patient-voice oriented condition, native Tagalog speakers mostly produced declaratives in Agent-Patient word order; however, in the agent voice, there were more Patient-Agent responses than Agent-Patient patterns. Tanaka concludes that while native Tagalog adults tended to produce patterns with a pivot-final word order, there was an observed tendency to still mention agents before patients. A similar study by Bondoc et al. (2018) was conducted with elderly native Tagalog adults, using only a set of patient voice-oriented items. It showed similar results: a preference for an Agent-Patient word order in patient voice. A concern with both studies, however, is that they involved limited numbers of participants and items, restricting generalization.

Garcia et al. (2018) investigated the word order tendencies of native Tagalog children and adults in declarative clauses using an oral sentence completion task. The items all involved pictures depicting reversible events. Participants heard a verb with voice morphology in either agent or patient voice and were instructed to complete the sentence based on what was depicted in the picture. In the agent voice, participants produced comparable numbers of Agent-Patient and Patient-Agent orders; in the patient voice, participants predominantly produced the Agent-Patient word order.

To summarize, the small number of psycholinguistic studies on Tagalog declarative sentences among native adults, as well as the theoretical claims on Tagalog, have revealed somewhat variable results on word order preferences, especially for the agent voice. Several questions remain, such as the word order tendencies in sentences with more than two arguments, as well as in understudied voice alternations, such as the benefactive and instrumental voice.
1.3.2. On Online Processing

This section reviews findings from two studies that probed word order preferences in Tagalog, and how sentential word order is processed in real time by the production and comprehension systems.

A production study using eyetracking by Sauppe et al. (2013) explored the time course of sentence formulation in Tagalog sentences involving transitive events. In terms of word order preference, native Tagalog speakers predominantly produced a pivot-last word order in both AV and PV when describing depictions of transitive events. Eyetracking patterns during these productions demonstrated that prior to speech onset, there were more fixations to the agent element than the patient, for all subsequent choices for voice (AV versus PV) or word order. Native Tagalog speakers nevertheless tended to have more fixations to the agent element when producing the agent voice relative to when producing the patient voice, in this pre-speech time window. This difference in agent fixations between voice types corresponded to a hypothesized grammatical encoding phase during speech planning. This hypothesized phase was followed by incremental linear encoding of the lexical elements, as evidenced by the increased looks to these elements based on their order of mention. While these findings contribute to knowledge of sentence production in Tagalog, there are some methodological concerns for this study. One of these concerns involves the modeling of pivot-final constructions in the practice phase of the experiment, which may have influenced the production of that specific sentence pattern in the participants, leading to a predominant number of pivot-final sentences in their data. Argument animacy, a factor that may affect word order preferences, was not controlled.

In a subsequent study, Sauppe (2016) investigated the types of information on the verb that native Tagalog speakers use to anticipate upcoming input. The properties of Tagalog allow a test of whether listeners orient more to syntactic or semantic information in verb processing. As Sauppe argued, these information types do not conflate in Tagalog verbs, unlike in other languages. He predicted that if native Tagalog speakers used morphosyntactic information (i.e. voice morphology) on the verb to make predictions, comprehenders would differ in their fixations to the upcoming first argument. With the assumption of a pivot-last word order preference in Tagalog (Himmelmann, 2005), comprehenders would have anticipated and have increased looks to a patient argument upon hearing a verb in agent voice, and an agent argument upon hearing a verb.
in patient voice. The alternative prediction in Sauppe’s study was for native Tagalog adults to use *verbal semantics* in prediction. Considering the relevance of agents in event cognition, comprehenders would conceptualize of an event involving an agent instigating an act upon a patient. Using semantic knowledge or event cognition would result to comprehenders fixating at agent argument upon hearing the verb, regardless of voice. The study outcomes showed increased fixations to the agent argument at the verb region, congruent to the prediction of the use of verbal semantics in anticipatory processing.

In the same comprehension study, he described the eye gaze patterns as adapting to the incremental word order of the sentence as it unfolded. The results showed a difference in fixation patterns between the agents and the patients at the first argument region, depending on the voice, and increased anticipatory looks to the upcoming second argument prior to the unfolding signal. He interpreted the fixation patterns as illustrating the integration of all information before the second argument is heard. With these results, Sauppe argued the use of verbal semantics and the integration of all information before the second argument as two potential anticipatory processes occurring in Tagalog comprehension. While Sauppe’s (2016) comprehension study on Tagalog was ground-breaking, some additional factors could have been considered in the experiment, such as the need to norm all stimuli items prior to the experiment to ensure their naturalness, and the need to control argument animacy. Also, the study was restricted to two of the common sentence patterns, which limited our understanding about the real-time processing that occurs in less preferred patterns (albeit this decision was justified by the study design and predictions). In this dissertation, I build on Sauppe’s experiment by creating stimuli that are controlled for animacy, and extend his work to look at other understudied voice patterns, i.e., the benefactive and instrumental voice.

In addition, the contradictory findings regarding the use of voice information in Sauppe’s production and comprehension studies raises some interesting questions. In production, speakers showed a difference between fixations to the agent and patient as an effect of voice prior to speech onset, suggestive of a grammatical encoding phase where speakers plan for the voice morphology to be used in the sentence. In contrast, in comprehension, comprehenders utilized verbal semantics (versus morposyntactactic information) to anticipate for the pivot; anticipatory looks were only observed on the upcoming second argument. The lack of voice effects in the comprehension study
seemed counterintuitive, considering the prominence of voice in Tagalog syntax. Sauppe (2016) interpreted these conflicting findings as a possible asymmetry between the production and the comprehension systems, as well as agent saliency potentially playing a more influential role in comprehension than in production. Although this could be a plausible explanation, these interesting findings call for future research on the use of voice information in comprehension.

Using the visual world paradigm, Garcia and colleagues (2020) examined how native Tagalog adults (as well as children of different age groups) use their word order knowledge and nominal marking to determine the thematic roles of the arguments in Tagalog sentences. Two pictures depicting reversible transitive events were presented on the screen, and participants were asked to click on the picture corresponding to the event in the heard sentence. Eye movement patterns were tracked to determine the time course of using nominal marking to determine the target picture of the sentence. Sentences included two voice patterns (AV and PV), crossed with two word order patterns (Verb-Agent-Patient and Verb-Patient-Agent), creating four conditions. Native Tagalog adults showed overall high accuracy scores across all conditions, and they were reported to use the nominal marking on the first noun phrase to disambiguate and look at the correct picture in all four conditions. While this study is also relevant, it did not explore whether there were effects of using voice morphology prior to encountering the noun argument, or whether voice morphology can be used to predict the likely upcoming arguments in the sentence in real-time.

1.4. Composition of the Dissertation

This dissertation aims to expand our understanding of prediction and expectation-based syntactic parsing by investigating whether the parser generates syntactic expectations that specify the relevant nominal arguments in particular sentence positions, such as the first and second nominal arguments after the verb. In addition, it investigates the hypothesis that voice morphology signals relevant information on word order that the parser immediately utilizes in a probabilistic manner to develop syntactic expectations when anticipating upcoming input. Tagalog provides an opportunity to examine this problem in prediction, given that the verbal morphology has critical implications for its sentential word order. Addressing this central question allows us to explore the ability of the parser to generate gradient predictions of syntactic structures beyond the word level, and to evaluate the immediate and fast-acting nature of prediction in sentence processing.
The following are the chapters present in this dissertation, as well as the experiments and/or discussion present in each.

In Chapter 2, I examine some of the grammatical factors that potentially influence linear word order in Tagalog, and probe whether these grammatical factors influence the anticipatory looks of native Tagalog speakers in real-time as they encounter the verb using the visual world paradigm. Experiment 1 confirms whether certain probabilistic grammatical constraints driving word order preferences could be deduced from certain types of linguistic information, i.e., agentivity and pivothood (Kroeger, 1993; Riesberg, 2014; Schachter, 1976). This experiment validates whether voice morphology could be a predictor of word order in Tagalog, a cue that listeners or readers may use when developing expectations about the input. I propose two probabilistic grammatical constraints, agent-first and pivot-second, as playing a role in linear word order preferences. An offline task (sentence completion) was implemented to determine the validity of these constraints. Experiment 2 investigates whether native Tagalog speakers use voice information to anticipate the likely nominal arguments in the unfolding input as soon as they encounter the verb. I hypothesize that the cognitive saliency of agents and pivots as two critical arguments in the syntactic representation will constrain the looks of Tagalog listeners to the agent and the pivot, as they hear the verb containing the voice morphology. This hypothesis was tested using the visual world eyetracking paradigm.

In Chapter 3, I address the question of whether speakers develop probabilistic syntactic expectations in real-time about the particular word order of a sentence using a series of reading experiments. Experiment 3 hypothesizes that speakers use voice information to generate probabilistic expectations relative to their word order preferences. These probabilistic expectations are hypothesized to be demonstrated in graded surprisal values, as they encounter the first morphological marking after the verb region. As enough information is accrued in the first NP region, the speakers’ beliefs are predicted to be updated, which will manifest as a lack of surprisal in the second NP region. To test these hypotheses, a self-paced reading task was implemented.

Additionally, this dissertation evaluates the notion that prediction is a fundamental principle and a routine process of the cognitive system. To assess the fundamentality of prediction as a cognitive process, I test whether speakers generate expectations in cognitively constraining contexts, such as in conditions with time pressure, or in conditions with increased memory load. I
hypothesize that the comprehenders’ expectations are modulated in cognitively constrained contexts. Experiments 4 and 5 examine the hypothesis that prediction, being one of the core operations of the human parser, results in native Tagalog listeners developing expectations about the linear word order of the sentence. To test the expectations of speakers, Experiment 4 involves the presentation of sentence stimuli using rapid serial visual presentation to induce time pressure. To test how these syntactic expectations change in real-time in cognitively constrained contexts, a self-paced reading task is conducted in Experiment 5, with the addition of participants producing a nonsense syllable to induce memory load.

Chapter 4 revisits the main question addressed by this dissertation on predictive processing, and discusses how the different dissertation experiments contribute to our current understanding of predictive processing.

Chapter 5 concludes the dissertation.
Chapter 2
Constraint Influences on Linear Word Order in Production and
Anticipatory Gaze Patterns in Comprehension

This chapter establishes the relationship between voice morphology and word order preferences in Tagalog. I examine the role of two factors, namely agent prominence and pivothood, in sentence production and comprehension. Two experiments are presented, which investigate the role of these two factors in ordering arguments in production, and in driving anticipatory gaze patterns in comprehension.

2.1. Experiment 1: Sentence Continuation Task

2.1.1. Introduction

The introductory chapter describes the variability of results in the small number of psycholinguistic studies on Tagalog declarative sentences among native adults, as well as the lack of consensus among scholars on the theoretical claims on Tagalog word order preferences. Researchers have claimed either an agent-first tendency, a pivot-last tendency, or that both tendencies influence linear word order in Tagalog. Several questions remain, such as what the word order tendencies are in sentences with more than two arguments, as well as in understudied voice alternations, such as the benefactive and instrumental voice.

Three factors are likely to influence the formulation of linear word order in Tagalog: pivothood or syntactic prominence, agent prominence, and animacy. These factors are briefly reviewed below.

2.1.1.1. Pivothood or Syntactic Prominence

As extensively reviewed in Chapter 1, The pivot element holds syntactic prominence in languages that exhibit Philippine-type voice. By syntactic “prominence”, I generally refer here to the eligibility of the pivot to participate in syntactic phenomena such as pseudo-cleft wh-questions (Schachter & Otanes, 1972), quantifier float, raising, and secondary predication (Kroeger, 1993;
Schachter, 1976), and relativization and control operations (Bondoc, 2018; Kroeger, 1993; Pizarro-Guevara & Wagers, 2018; Schachter, 1976). Other perspectives have also suggested the pivot as a marker of information structure (Chen, 2017; Rackowski & Richards, 2005).

2.1.1.2. Agent Prominence

There is an abundant literature in formal linguistic work that has shown agents to be less marked and more highly ranked in hierarchies compared to other thematic roles (Aissen, 1999; Croft, 1991; Langacker, 1995). Some Tagalog verbal morphology that expresses volition and ability incorporates agent properties, highlighting the prominence of agents (Riesberg & Primus, 2015).

In production, speakers find it easier to produce sentences with agents mapped as sentential subjects, in languages that have a subject-first word order (Ferreira, 1994; Kemmerer, 2012), although this is highly mediated by factors such as animacy, as reviewed below.

In comprehension, classic psycholinguistic work described a tendency for the system to initially parse the first noun as the agent, resulting in a processing advantage in sentences that begin with an agent (Bever, 1970; Townsend & Bever, 2001). The notion of agent prominence has also gained support from research that showed comprehenders process agents implicitly (Mauner et al., 1995; Mauner & Koenig, 2000). In “good-enough” processing models, comprehenders tend to apply heuristic strategies that allow them to identify the first nominal as the agent of the event (Ferreira, 2003; Ferreira et al., 2002, 2009; Karimi & Ferreira, 2016). In psycholinguistic work on Tagalog, a preference for agent-first word orders has been observed in multiple populations, such as adults with aphasia (Bondoc et al., 2018) and typically-developing children (Bautista, 1983; Garcia, Dery, et al., 2018; Garcia, Roeser, et al., 2018; Garcia et al., 2020; Segalowitz & Galang, 1978; N. Tanaka, 2016).

Agent prominence in the processing of events is also well-recognized in neurocognitive research (Bornkessel-Schlesewsky & Schlesewsky, 2009, 2013; Sauppe, 2016; Sauppe et al., 2013), and more general psycholinguistic work (Cohn & Paczynski, 2013; Ferreira, 2003; Segalowitz, 1982). For example, research from gesture and pantomime studies has shown a preference to sequence arguments depending on how events unfold in the real world, with the
agent or the instigator of the action acting upon a patient or theme (Goldin-Meadow, 2005; Goldin-Meadow & Feldman, 1977; Goldin-Meadow, So, Özyürek, & Mylander, 2008).

2.1.1.3. Animacy

Because entities that are highly agentive also tend to be animate, there is a close link between agentivity and animacy, and animacy is also an influential constraint that can impact language production and comprehension. As will be discussed later, animacy will be one of the critical variables that I will manipulate in my experiments.

The influence of animacy in linguistic phenomena has been described in formal linguistic theories such as in animacy hierarchies (Comrie, 1989; de Swart et al., 2008; Silverstein, 1976; Yamamoto, 1999), and animacy-influenced thematic hierarchies (Evans, 1997; Givón, 1984). In production, animacy relates to conceptual accessibility: elements that are animate tend to be mentioned earlier in the sentence as they are highly accessible, and are likely to be mapped as the sentential subject (Branigan et al., 2008; Christianson & Ferreira, 2005; Ferreira, 1994; McDonald et al., 1993; M. N. Tanaka et al., 2011). In comprehension, animacy bears on the initial interpretation of syntactically ambiguous utterances (Clifton Jr et al., 2003; Ferreira & Clifton Jr, 1986; Trueswell et al., 1994). A general cognitive advantage for animates is also reported in the literature, as they easily capture and sustain more visual attention compared to inanimates (Abrams & Christ, 2003; Calvillo & Hawkins, 2016; Guerrero & Calvillo, 2016; Pratt et al., 2010) and lexical items represented by animates are better remembered than those by inanimates (Bonin et al., 2014; Leding, 2019; Nairne et al., 2013; Nairne & Pandeirada, 2010; VanArsdall et al., 2015).

In neurocognitive research, animates and inanimates have been reported to be processed in two different parts of the brain (Caramazza & Shelton, 1998; Sha et al., 2015). The match or mismatch of the animacy features of an unfolding argument can affect event-related potentials (ERPs), suggesting that animacy is one of the constraints that the mind computes as it develops expectations for upcoming input (A. Kim & Osterhout, 2005; Paczynski & Kuperberg, 2011, 2012; Szewczyk & Schriefers, 2011, 2013; Wang et al., 2020). Modeling work has shown that animate entities share more common features than inanimates, entailing a difference in the taxonomical structure between these, and a possible advantage with processing animates than inanimates (Kemp & Tenenbaum, 2008; Rogers & McClelland, 2008).
2.1.2. Hypotheses

Theoretical debates on word order, along with the limited experimental work on this topic, call for more data on word order preferences in Tagalog, as well as further investigation of the factors that drive these preferences. The correct generalizations about word order remain unclear, especially for extensions beyond simple transitive sentences in AV or PV. For example, how strong is the tendency to place the agent versus the patient – the two core arguments – in the first nominal position when neither of these is the pivot? Does the animacy of the pivot affect syntactic prominence? I also examine the influence of animacy in evaluating the relative strength of pivot prominence versus agent prominence by extending to four different voices that involve pivots that are either animate (agent and benefactive voice) or inanimate (patient and instrument voice).

Several factors can influence word order choices, including conceptual accessibility and syntactic properties. Taking an information processing perspective, and aligning with earlier descriptions of Tagalog word order, I hypothesize a processing preference for relevant elements (e.g., dependents) to be closer to heads (e.g., verbs) when permitted by the grammar, as it promotes their accessibility for syntactic computations (Gibson, 2000; Hawkins, 2004; Jaeger & Tily, 2011; Lewis et al., 2006; MacDonald, 2013). In Tagalog, for which the grammar licenses any word order, I construe agents and pivots as prominent elements that the processing system would track, for reasons that I will explain shortly, and preferentially place close to the verb.

The existence of more than one prominent element in Tagalog may result in a competition for these elements to be in an argument position closest to the verb. Variation could thus exist: one could describe the likelihood of agents or pivots to be in the first argument position. I hypothesize that if either the agent or the pivot fills in the first position, the other element is likely to occupy the second position. Furthermore, other, less relevant elements may also be placed in these positions close to the verb, as the flexible word order allows. There are thus multiple word order permutations, which could be described in terms of probabilities, such as a probability for agents to be first, agents to be second, pivots to be first, pivots to be second, and so on. Within a multiple constraints system, it could also be imagined that decreasing the animacy of the pivot could weaken the tendency of the pivot to be closer to the verb. The following are two constraints I hypothesize
to be important to Tagalog word order. I then offer a hypothesis that addresses the relative weight or strength of the two tendencies for situations in which they conflict.

**Hypothesis 1. Agent-first constraint:** Agents, when present in a sentence, tend to be the first post-verbal nominal argument in Tagalog.

If we assume that the grammar allows multiple word orders, a tendency for agent-first orders could be rooted in agent prominence within the cognitive system. As mentioned in the Introduction, a variety of perspectives suggest increased attention to the agent argument compared to other arguments by the language processing system, as well as a tendency to place agents early in the sentence, owing to their conceptual accessibility (e.g., Cohn & Paczynski, 2013; M. N. Tanaka et al., 2011).

**Hypothesis 2. Pivot-second constraint:** Pivots tend to occur in the second argument position in Tagalog.

The structural configuration of Tagalog privileges the pivot element to hold syntactic prominence. Other scholars have also suggested the pivot as a marker of information structure (Chen, 2017; Rackowski & Richards, 2005), adding to its prominent status. This central role of the pivot in Tagalog syntax may then induce native Tagalog speakers to track this element in syntactic computation and place it near the verb. Because the grammar allows flexible word order, pivots are preferentially placed in the second argument position, which has computational advantages over later positions.

The simplest version of the Pivot-second hypothesis maintains that the strength of pivot prominence will not differ, regardless of the animacy of the pivot. Thus, in voices that have inanimate pivots (patient and instrument voice, in the experiments below), I would also expect the pivot-second constraint to apply.

I take this constraint to apply to all pivots, including agent pivots. That is, if an argument functions as the pivot, this constraint pushes it to be placed in the second position. An additional prediction is also made for the agent voice, then, where these two constraints conflict.
Hypothesis 3. The agent-first constraint is more heavily weighted than the pivot-second constraint:

I hypothesize that the constructs of agent and pivot prominence also vary in weight, and that agent prominence is more heavily weighted than pivot prominence, due to its general cognitive saliency, as reviewed above. As a result, I hypothesize that the preference for agents to placed in the position closest to the verb (first argument position) would dominate a preference to place pivots in the first position, for all voices – leading to the development and maintenance in the language of the pivot-second preference discussed above. For the agent voice, I hypothesize that the agent-first tendency would dominate (although not completely eliminate) the influence of the pivot-second tendency. This would result in native Tagalog speakers exhibiting some degree of preference for an Agent-Patient-Oblique linear word order in the agent voice.

2.1.3. Method
The study hypotheses were experimentally tested using a sentence continuation task. This task also served as a norming study for Experiment 2. Participants were asked to complete a sentence fragment by mentioning all four entities presented in a 2 x 2 visual array, the complete details of which will be presented below.

2.1.3.1. Participants
Thirty adults (age range 22-55, mean age 29.8; 14 female and 16 male) living in Metro Manila participated in the study. All participants identified themselves as native Tagalog speakers and as Tagalog-dominant in a language background questionnaire. All participants also spoke English as a second language. Some participants also identified themselves as a speaker of a third language. Overall, participants expressed varying degrees of proficiency for their second and third languages.

2.1.3.2. Materials
The original norming study involved 52 verbs, taken from a Tagalog verb subcategorization study (Endriga, 2014). These verbs belong to a class that could take an agent, a patient, a benefactive, and an instrument as a pivot. The affixal forms were also controlled for each voice, such that all verbs take [nag-] for agent voice, null affix for patient voice, [ipag-] for benefactive voice, and
[ipang-] for instrumental voice. Four presumed prototypical arguments (an agent, a patient, a benefactive, and an instrument) were also constructed for each verb; the animacy of the arguments was controlled by having two animates (agents and benefactives) and two inanimates (patients and instruments).

The 52 verbs were crossed with four voice alternations (agent, patient, benefactive, and instrument voice) in a Latin-square design, and with two visual scene arrangements, creating eight lists. To avoid the subsequent occurrence of two items of the same voice pattern, all items in each list were pseudo-randomized. Participants were assigned in a rotating fashion to one of these eight lists.

In all pictures, the animates appeared in the top two positions of the 2 x 2 visual array and the inanimates were positioned in the bottom two positions. Two sets of pictures were created for each verb: in one set, the agents and patients were positioned on the left, and the benefactives and instruments were on the right. In the other set, the positions of these elements were switched. Note that these pictures were also prepared and normed for the visual scenes in Experiment 2.

Because it was anticipated that items might vary in how well participants recognized the intended arguments from their pictures, recognized the intended thematic roles, agreed with each other on the labels for the arguments, and found the item to be natural, all items were evaluated for these features before retaining them for the analysis of word order presented in the results section. The accuracy of participants in providing the target responses of the sentence continuation task were scored using binary coding. Each argument in the sentence was scored in terms of recognizability from the visual stimulus, correct thematic role assignment, and correct use of syntactic markers. An overall score was computed for each item. These scores were then transformed to z-scores per participant, and the z-score values were then aggregated by verb and voice. A similar z-score transformation measure was made from the overall sentence ratings. Verbs that were beyond 1.5 standard deviations from the mean score in at least one voice type were excluded. The final set of items from this process involved 37 verbs, and only the results from these 37 verbs were included.9

9Similar word order patterns were found in the full set of items. The stimuli items used for Experiment 1 are available in Appendix C. All data, R scripts including statistical analyses, and other stimuli materials in the experiment may be accessed on my OSF account (https://osf.io/n5tjy/).
2.1.3.3. Procedure
The experiment was implemented via Google Survey Forms. Native Tagalog speakers were recruited from the Philippines via advertisements and word-of-mouth. Each sentence fragment had a verb with the appropriate voice morphology for the condition, and a following adverbial phrase (Figure 1). Participants completed the sentence fragment by typing their responses into a textbox. Participants were asked to use all four entities that were depicted in the visual scene, with each element to be used only once. They were also asked to evaluate the naturalness of the sentence they completed using a 6-point scale, with 1 being ‘very unnatural’ and 6 being ‘very natural.’

Figure 1. A sample item in the online sentence completion task. The visual stimulus displays the four elements mother, child, spaghetti, and pan for the verb cook. The test sentence is ‘Ipinandurog kani-kanina lang...’, which roughly translates to ‘IV-mash a while ago...’ See sentences (3) to (6) in Chapter 1 for sample completions.

2.1.3.4. Analysis
Participant responses were coded in terms of the linear order of the arguments in the sentence. All arguments were assessed in terms of their correctness or grammatical acceptability. To ensure coding reliability, a second coder, who is a native Tagalog speaker and has knowledge of linguistics, was asked to code the same data. There was 89.9% agreement and $\kappa = 0.73$ on all analyzed trials, suggesting a substantial level of agreement between the two coders (Cohen, 1968). The guidelines for the evaluation of responses are included in Appendix A.
Word order analyses were segmented in terms of argument position. As an example, sentence (2) would be coded as AG-PAT-BEN-INS for linear word order. This coding would reflect that an agent was used in the first argument position, a patient was used in the second argument position, a benefactive in the third, and an instrument in the fourth position. Guidelines for syntactic coding are provided in Appendix B.

To test for the validity of the constraints, the distributions of nominal elements mentioned in the first and second argument positions were determined. Two binary measures were created to test the study hypotheses. The agent-first constraint was tested by scoring responses that had an agent in the first position with a “1”; otherwise, a “0” was given. Likewise, a similar binary coding measure was adopted for the pivot-second constraint. The likelihood of the agents (versus other) mentioned in the first position, and the pivots (versus other) mentioned in the second position, were tested using maximal mixed effects logistic regression models (Bates et al., 2015), with Voice as a fixed effect, and the intercepts and slopes of the fixed effects included for participants and items as random effects, using R (R Core Team, 2013). Due to non-convergence, the predictor Voice was nested in the random effect for Item in one model.

The predictor Voice was contrast coded in all models using dummy coding, with the agent voice set as the reference level for all models, to assess whether the claims on agent voice patterns were supported in this condition. Dummy coding also allowed a comparison of agent voice with the other voice patterns to determine an effect of Voice.

2.1.4. Results

An overall assessment of the correctness (or grammatical acceptability) of responses in the sentence continuation task showed the participants committing more erroneous responses in the benefactive and instrumental voices than in the agent and patient voices (Figure 2).

Further analysis of scores showed that the large number of errors in the benefactive and instrument voice patterns were almost evenly distributed across participants and items. I therefore attribute the higher error rates in the benefactive and instrumental voices to potential factors such as their relatively lower frequency, higher processing difficulty, or need for a more constrained discourse context, and leave further investigation of the cause to future research.
2.1.4.1. Correct Responses

![Figure 2](image)

**Figure 2.** The number of responses with correctly and incorrectly marked arguments for each voice type.

After the exclusion of the incorrect responses, the frequencies of the different types of arguments mentioned in the first and second position were analyzed (Figure 3; see also Table 4 in the Discussion section). In the first position, for the agent voice, mentions of agents slightly exceeded those of patients, while other arguments were quite rare. In the first position for the other voice types, mentions of agents substantially exceeded mentions of patients or other argument types. Relatively small numbers of adjuncts such as benefactives and instruments were mentioned in the first position for the agent and patient voices (see Table 4). These adjunct counts slightly increased in the benefactive and instrument voices.

In the second position, the pivot element was the predominant element mentioned in all voice types except the agent voice, where more patients were mentioned compared to agents. The pivot element prevailed in the second position regardless of the animacy of the pivot: Notice the similarity in pivot-second patterns for the BV and IV. Notably, though, other elements were also produced in the second position. Non-agent, non-patient elements comprised 12.65% of the second-position mentions for the agent voice; non-pivot elements comprised 12.89 to 35.56% of the second-position mentions for the other voice patterns.

These findings suggest that statements about agent and pivot placement should be tempered in some fashion. It is certainly not the case that these productions are strictly agent-first, pivot-

---

10The increase in the mention of benefactives and instruments in the first position in the BV and IV may reflect the tendency to immediately mention the pivot after the verb. It is critical to point out, regardless, that in these two voice patterns, the agent remains the predominantly mentioned element in the first position.
second, or pivot-last. Despite the flexible word order of the language, animacy does not appear to play an important role in determining which argument appears in the first or second position. However, the patient argument, which was inanimate in all of the experiment materials, was often placed in the first or second position. I return to these points in the Discussion section.

Figure 3. Frequency of the different arguments in the first (left panel) and second position (right panel) in correct responses. Arguments marked with a cross indicate targeted argument responses.

The outcome of the mixed effects logistic regression model that tested the agent-first hypothesis (Table 2) showed a non-significant intercept value in the agent voice, indicating that the numerical preference for agents in the first position in the agent voice was not significant ($p = 0.61$). That is, the preference for agents was not significantly higher than 50%, when the data are treated as a choice between agent and non-agent. However, a statistically significant difference between agent voice and each of the other three voice patterns, i.e., the patient, benefactive, and instrumental voices, was observed, indicating that each of the other voices increased mention of the agent in the first position ($p < .001$). The preference for the agent to be in the first position was strong regardless of the animacy of the pivot.

The second model, testing for the pivot-second hypothesis, resulted in a statistically significant intercept with a negative estimated $b$ value. This result indicated a significant dispreference for the agent to occur in the second position in agent voice ($p < .01$), when it functions as the pivot, versus
any other element. In addition, statistically significant differences were observed between the agent and patient voice \((p < .001)\), the agent and benefactive voice \((p < .01)\), and the agent and instrumental voice \((p < .001)\), indicating that each of the other voices increased mention of the expected pivot in the second position. Importantly, with the exception of agent voice, pivots were strongly preferred in the second position regardless of whether the pivot was animate (such as in benefactive voice) or inanimate (such as in patient and instrument voice). I further discuss the implications of these findings in the Discussion section.

Table 2. Results of the mixed effects logistic regression models that tested the likelihood of the agents being the first argument, and the pivots being the second argument.\(^{11}\)

<table>
<thead>
<tr>
<th></th>
<th>Agent-first hypothesis</th>
<th>Pivot-second hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(\beta)</strong></td>
<td><strong>SE</strong></td>
<td><strong>z</strong></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.14</td>
<td>0.26</td>
</tr>
<tr>
<td>AV vs PV</td>
<td><strong>2.94</strong></td>
<td><strong>0.55</strong></td>
</tr>
<tr>
<td>AV vs BV</td>
<td><strong>1.40</strong></td>
<td><strong>0.37</strong></td>
</tr>
<tr>
<td>AV vs IV</td>
<td><strong>1.88</strong></td>
<td><strong>0.48</strong></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.66</td>
<td><strong>0.24</strong></td>
</tr>
<tr>
<td>AV vs PV</td>
<td><strong>2.93</strong></td>
<td><strong>0.43</strong></td>
</tr>
<tr>
<td>AV vs BV</td>
<td><strong>1.37</strong></td>
<td><strong>0.44</strong></td>
</tr>
<tr>
<td>AV vs IV</td>
<td><strong>2.02</strong></td>
<td><strong>0.45</strong></td>
</tr>
</tbody>
</table>

While the predominance of agents and pivots in the predicted positions for the patient, benefactive, and instrumental voice patterns was consistent with the study hypotheses, the almost equal mention of agents and patients in the first and second positions in the agent voice called for additional analysis. The earlier binary scoring only considered the likelihood of one type of element occurring in a given position, disregarding the other types of arguments that could precede or follow it. Hence, the non-preference for agents to occur in either the first or second position in agent voice was further probed through a subset analysis where only responses with agents and patients

\(^{11}\)Mixed effects logistic regression models used
Agent-first hypothesis: glmer(testagentfirst ~ Voice + (1+Voice|Participant) + (1+Voice|Item), data, family="binomial")
Pivot-second hypothesis: glmer(testpivotsec ~ Voice + (1+Voice|Participant) + (1+Voice|Item), data, family="binomial")
occurring in the first two positions were considered, since these are the core arguments and the two nominals most likely to occur in the first two positions. The results of the mixed effects logistic regression models (Table 3) demonstrated a non-significant value for the intercept for both the agent-first hypothesis and the pivot-second hypothesis, indicating a non-significant likelihood for agents to occur in either the first or second position, i.e. the lack of evidence for a preference between agents and patients in either position. It is critical to point out, however, that there was an animacy difference between agents and patients, which might have enhanced the preference for agent-first order (and thus the agent-patient order in AV). Unsurprisingly, a significant difference between agent and patient voice was observed in each pairwise comparison, confirming the effects of patient voice shown in the primary analysis ($p < .001$ in the agent-first hypothesis; $p < .01$ in the pivot-second hypothesis).

Table 3. A subset analysis of the responses with either agents or patients in the first and second positions in the agent voice and patient voice conditions, using mixed effects logistic regression. 

<table>
<thead>
<tr>
<th></th>
<th>Agent-first hypothesis</th>
<th>Pivot-second hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>$SE$</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.33</td>
<td>0.27</td>
</tr>
<tr>
<td>AV vs PV</td>
<td>27.08</td>
<td>5.72</td>
</tr>
</tbody>
</table>

12 Since the only aim of the subset analysis was to determine the likelihood of the agent versus the patient to occur in the first two positions in the agent voice, only one condition (the agent voice) was of interest. However, having at least one other condition (patient voice) for comparison allowed the retention of the predictor Voice in the model. The two other voice conditions were not included in this data subset.

13 Mixed effects logistic regression models used
   Agent-first hypothesis: glmer(testagentfirst ~ Voice + (1+Voice|Participant) + (1+Voice|Item), data, family="binomial")
   Pivot-second hypothesis: glmer(testpivotsec ~ Voice + (1+Voice|Participant) + (1|Item) + (1|Item:Voice), data, family="binomial")
2.1.4.2. Incorrect Responses

Due to the numerous errors committed by the participants in the experiment in the benefactive and instrumental voices, the distribution of arguments in the incorrect responses was also examined. There remained a predominant production of the agent argument in the first position in these responses, and of the *ang*-marked element (regardless of an incorrect pivot selection) in the second position (Figure 4). These findings provide additional support for the agent-first and pivot-second hypotheses. I leave further analysis of these errors to future research.

![Figure 4](image)

**Figure 4.** Frequency of the different arguments in the first (left panel) and second position (right panel) in incorrect responses. Arguments marked with a cross indicate targeted argument responses. Arguments marked with an asterisk indicate incorrect forms (*ang-X denotes an incorrectly selected element as the pivot; *X-Argument denotes arguments with incorrect nominal marking).

2.1.5. Discussion

Experiment 1 examined the extent to which agent and pivot prominence influences production by investigating how these factors impact word order preferences in a language with flexible word order, Tagalog. The results obtained from the mixed-effects regression models for the full dataset strongly support the applicability of the hypothesized agent-first and pivot-second constraints in the patient, benefactive, and instrumental voices, highlighting the influences of agent and pivot prominence to linear word order. In the agent voice, comparable numbers of agent and patient
arguments were mentioned within the first and second positions. The lack of a significant preference for agents in the first position and the significant dispreference for agents in the second position (compared to all instantiated nominals) provide only partial support for the prediction that an Agent-Patient order would dominate in AV. Instead, they suggest that the agent can freely occur in either position. A subset analysis aimed to verify this finding by considering only the responses where the agent and the patient filled the first two positions (i.e., Agent-Patient or Patient-Agent as the first two arguments). The results failed to find a preference for the agent relative to the patient in either the first or second position of AV, implying that in the agent voice the Agent-Patient and Patient-Agent linear word order patterns were similarly accessible, at least for the types of sentences tested here. This finding is in consonance with Garcia et al. (2018) and Hsieh (2016), and partly with Kroeger (1993). Importantly, this outcome entails that the relative strengths of agent and pivot prominence with regard to linear word order are comparable, contradicting the third hypothesis in this experiment.

It is also relevant to underscore that animacy did not appear to affect pivot placement; with the exception of agent voice, the pivot element dominated the second position in voice patterns involving an animate (BV) or an inanimate pivot (PV, IV). This result suggests that pivot prominence overrides the influence of animacy in Tagalog production. This finding runs against incremental, lexically-driven models of sentence production that propose that conceptual accessibility overrides syntactic function, with animacy being one of those factors related to conceptual accessibility (De Smedt, 1990; Gleitman et al., 2007; Kempen & Hoenkamp, 1987), and supports hierarchically incremental models that first formulate the event representation and consider the thematic role and syntactic function assignment of the lexical items (Bock & Irwin, 1980; Bock & Warren, 1985a; Lee et al., 2013). Hierarchically incremental models of production would be feasible for Tagalog, given that the verb-initial syntax may require speakers to first plan the grammatical voice, i.e., the relationship between the verb and the syntactically prominent element. Similar support for hierarchical models for verb-initial languages such as Tagalog has been shown by Sauppe et al. (2013), as well as for other understudied languages that have flexible word order such as Tzeltal (Norcliffe et al., 2015), Odawa/Ottawa (Christianson & Ferreira, 2005), and Murrinhpatha (Kidd et al., 2018). It remains to be investigated in future research whether the
relative strength of animacy versus syntactic prominence such as pivothood could vary depending on the syntactic phenomena and discourse context.

While agents and pivots were the most frequently mentioned elements in the first and second positions, the patient also received numerous mentions in either the first or second position. It can be argued that the patient element also carries some degree of prominence, albeit less than that of agents and pivots, as one of the core arguments (Himmelmann, 2005; Riesberg, 2014). The status of an element as either core or oblique reflects grammatical relations as well. The core/oblique contrast relates to one of the syntactic relationships that speakers compute as they plan for the event during message formulation, aligning with hierarchical incremental models. Future research may also further examine the critical role of a core/oblique constraint in production.

The following patterns summarize the preferred word order patterns in Tagalog; the pivot argument is underlined:

(i) AV \{Verb + Agent + Patient + \ldots; Verb + Patient + Agent + \ldots\}
(ii) PV Verb + Agent + Patient + \ldots
(iii) BV Verb + Agent + Benefactive + \ldots
(iv) IV Verb + Agent + Instrument + \ldots

The evidence that cognitive constraints such as agent and pivot prominence influence syntactic linear word order patterns helps illuminate how the cognitive system responds to production demands. From the perspective of processing, because of the cognitive system’s resource limitations (e.g., in memory or attention), activating highly accessible information and syntactically relevant elements and mentioning them earlier eases the burden on sentence formulation (Gibson, 2000; Hawkins, 2004; Jaeger & Tily, 2011; Lewis et al., 2006; MacDonald, 2013). From the perspective of constraints and information sources, word order patterns provide evidence for the critical role of agents and pivots in the syntax of the language.

The interaction of these two cognitive constraints with other factors such as a core/oblique constraint highlights the importance of a multiple-constraints system in explaining production outcomes. One source of evidence for the advantage of the multiple constraints approach in describing production performance is the range of word order distributions across voices. While there were dominant word order patterns for each voice, there were also smaller counts of other elements mentioned in the first and second argument positions, as displayed in Table 4. These
patterns are grammatically acceptable, despite their lower frequencies, and emerged even though factors such as animacy and the discourse context were controlled in the experiment. A multiple constraints approach allows us to capture finer-grained details of production such as these subordinate word order patterns.

Finally, the present findings on Tagalog word order preferences bear on syntactic theory. Representing Tagalog with a single, canonical word order (plus movement or scrambling operations), similar to what has often been done with more frequently studied Indo-European languages may not be sufficient, as such operations only capture the flexibility of word order, and not the attested proportions of patterns in the language. Both the predominance and the variability of syntactic patterns must be covered in a complete account of word order in Tagalog, as well as in other flexible word order languages. Currently, the dominant theories in syntax only offer rules and principles that provide an explanation for a language’s flexible word order, and presume processing explanations for the attested frequencies or probabilities in the language. Future research must either reframe current frameworks of syntax to incorporate both flexibility and dominance, pin down the role of other factors such as discourse and information structure in word order, or extend research on processing to shed light on the interaction of grammar and processing in explaining the complex phenomena of Tagalog word order.

Having established that the agent and pivot are prominent in each of the voices I tested, leading to identifiable word order patterns for each voice, I now turn to the comprehension experiment. In Experiment 2, I ask whether the prominence patterns I found in production are mirrored in comprehension.

Table 4. Counts of elements appearing in the first and second argument positions.

<table>
<thead>
<tr>
<th>Voice</th>
<th>Elements</th>
<th>n</th>
<th>%</th>
<th>Voice</th>
<th>Elements</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV</td>
<td>Agent</td>
<td>142</td>
<td>52.79</td>
<td>AV</td>
<td>Agent</td>
<td>97</td>
<td>37.16</td>
</tr>
<tr>
<td></td>
<td>Patient</td>
<td>125</td>
<td>46.47</td>
<td></td>
<td>Patient</td>
<td>131</td>
<td>50.19</td>
</tr>
<tr>
<td></td>
<td>Instrument</td>
<td>0</td>
<td>0</td>
<td>Instrument</td>
<td>12</td>
<td>4.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genitive</td>
<td>0</td>
<td>0</td>
<td>Genitive</td>
<td>12</td>
<td>4.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Locative</td>
<td>2</td>
<td>0.74</td>
<td>Locative</td>
<td>8</td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coordinated</td>
<td>0</td>
<td>0</td>
<td>Coordinated</td>
<td>1</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>PV</td>
<td>Agent</td>
<td>238</td>
<td>91.89</td>
<td>PV</td>
<td>Agent</td>
<td>8</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>Patient</td>
<td>20</td>
<td>7.72</td>
<td></td>
<td>Patient</td>
<td>223</td>
<td>87.11</td>
</tr>
<tr>
<td></td>
<td>Instrument</td>
<td>1</td>
<td>0.39</td>
<td>Instrument</td>
<td>15</td>
<td>5.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genitive</td>
<td>0</td>
<td>0</td>
<td>Genitive</td>
<td>3</td>
<td>1.17</td>
<td></td>
</tr>
</tbody>
</table>
2.2. Experiment 2: Visual World Eyetracking

2.2.1. Introduction and Hypotheses

Experiment 1 demonstrated that agent and pivot prominence influence linear word order in production, as described by the agent-first and pivot-second constraints. In Experiment 2, I examine whether agent and pivot prominence drive anticipatory gaze patterns in comprehension. The word order patterns in Experiment 1 suggest that voice morphology in itself can serve as a cue that comprehenders can use to anticipate the agent and the pivot as the two arguments most likely to follow the verb. Hence, I predicted that agent and pivot prominence would drive anticipatory looks to both the agent and the pivot, resulting in more looks to these two entities than to depictions of the remaining arguments in the time period prior to the mention of the first nominal argument. Such a pattern follows from models that suggest that the production system is employed in comprehension to generate predictions of upcoming material (Dell & Chang, 2013; MacDonald, 2013; Pickering & Gambi, 2018; Pickering & Garrod, 2007, 2013), and previous work within the visual world paradigm that has shown anticipatory looks to images consistent with predicted sentential material (e.g., Altmann & Kamide, 1999; Kaiser & Trueswell, 2004; Kamide et al., 2003; Knoeferle et al., 2005).
Although motivated by theoretical claims about prediction, the limited psycholinguistic work on Tagalog has not demonstrated the use of grammatical information in the form of voice morphology or pivothood in the anticipatory looks; only an effect of agent prominence was observed in previous work (Sauppe, 2016). The experiment presented here provides more opportunity for effects of pivot prominence to emerge, by making use of carefully normed stimuli that include additional, less prominent arguments to compare to the agent and pivot arguments, and by incorporating an animate non-agent entity as the pivot. The following specific hypotheses are tested in this experiment:

**Hypothesis 1. Increased looks to the pivot:** Comprehenders will use the voice information present in the verb to pre-activate the pivot as the most syntactically prominent argument of the verb and an argument that preferentially appears soon after the verb. This prediction will be manifested as increased looks to the pivot at the verb region, i.e. during and immediately following the mention of the verb.

More specifically, the proportion of looks to an argument X in X voice will be greater than the proportion of looks to argument X in any non-X voice. For instance, given a patient-voice verb, the proportion of looks to the patient is expected to be greater in patient voice than in non-patient voices (i.e., agent voice, benefactive voice, or instrumental voice).

<table>
<thead>
<tr>
<th>X</th>
<th>X Voice</th>
<th>X</th>
<th>Non-X Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent</td>
<td>in AV</td>
<td>&gt; Agent</td>
<td>in PV, BV, IV</td>
</tr>
<tr>
<td>Patient</td>
<td>in PV</td>
<td>&gt; Patient</td>
<td>in AV, BV, IV</td>
</tr>
<tr>
<td>Benefactive</td>
<td>in BV</td>
<td>&gt; Benefactive</td>
<td>in AV, PV, IV</td>
</tr>
<tr>
<td>Instrument</td>
<td>in IV</td>
<td>&gt; Instrument</td>
<td>in AV, PV, BV</td>
</tr>
</tbody>
</table>

**Hypothesis 2. Increased looks to the agent:** With agent prominence, and with the agent as one of the most probable immediately upcoming elements after the verb, agents will be pre-activated upon hearing the verb and voice morphology, resulting in increased looks to the agent in the verb region. Hence, there will be a strong tendency to look at the agent in all voices.

**Hypothesis 3. Looks to the agent will be greater than looks to the pivot:** I hypothesize that the strength of agent prominence will be greater than pivot prominence; thus, there will be more looks to the agent element than to the pivot element in each voice during the verb region.
Additionally, given that the agent is also the pivot in the agent voice condition, there will be an increased boost in activation in the agent argument in the agent voice, compared to the other conditions. This increased boost will result in the agent voice having the highest looks to the agent area of interest, compared to the other voice conditions, in the verb region.

2.2.2. Method

This experiment employed the visual world eyetracking paradigm. In this task, participants heard a sentence while looking at four entities presented in 2 x 2 visual arrays (Figure 1). Participant fixations to the different entities were recorded throughout the trial.

2.2.2.1. Participants

Forty-nine native Tagalog speakers (age 18-79, mean 36.54) living in Hawaii were recruited for the experiment through advertisements or word-of-mouth. All participants had normal (or corrected-to-normal) vision and normal hearing. All participants identified themselves as Tagalog-dominant in a language background questionnaire. All participants also spoke English as a second language. Some participants also identified themselves as speaking a third or fourth language, in varying degrees of proficiency. Thirty-four participants identified themselves as speakers of the Metro Manila Tagalog variant; 6 other participants spoke Metro Manila Tagalog and another Tagalog variant, while 9 spoke a non-Metro Manila Tagalog variant.

2.2.2.2. Materials

The visual scenes and linguistic stimuli were drawn from the materials normed by Experiment 1. Refer to sentences (3) to (6) in Chapter 1 for example sentence stimuli.

2.2.2.2.1. Critical Items

Thirty-six verbs from the final set of items in Experiment 1 were used in the eyetracking experiment. Similar controls for verb class and affixation were applied in this experiment. The most frequent argument labels from the participant responses in Experiment 1 were used as nominal arguments for the experimental sentences in Experiment 2. The postverbal adverbial expressions used in the sentence fragments in Experiment 1 (around 5-6 syllables in length) were
also used to expand the verb region and allow ample time to fixate the agent and the pivot in sequence.

These 36 verbs, which constituted the critical items, were crossed with the four voice conditions in a Latin-square design. These verbs were further crossed with the counterbalanced visual displays in Experiment 1 to create eight presentation lists.

All visual displays carried a high probability of eliciting the intended thematic role assignments of the agent and the pivot, as verified from the sentence completions in Experiment 1. Table 5 summarizes the percentages of intended agents/pivots used by the participants in their sentence completions (see the Discussion section for a more detailed analysis of these thematic roles).

<table>
<thead>
<tr>
<th></th>
<th>AV</th>
<th>PV</th>
<th>BV</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>% mention of intended agent</td>
<td>82.58%</td>
<td>83.07%</td>
<td>80.63%</td>
<td>80.95%</td>
</tr>
<tr>
<td>% mention of intended pivot</td>
<td>82.58%</td>
<td>91.73%</td>
<td>81.88%</td>
<td>94.56%</td>
</tr>
</tbody>
</table>

To preview the results, the study did not find an effect of voice on anticipatory looks; a main animacy effect was found instead. Hence, further analysis was conducted by examining whether a similar animacy effect was found for the fillers. In the next section, I briefly report on the nature of the filler items.

2.2.2.2.2. Filler Items

In each presentation list, the critical items were interspersed with 72 filler items which involved verbs of various types, such as intransitive unergatives, intransitive unaccusatives, locative voice sentences, and transitive sentences with experiencer pivots. The verbs used in the fillers varied in terms of voice morphology, and there was variation in the order of the nominal elements depending on the verb type and voice morphology. To maintain consistency across all experimental items, the filler items were similar in length to the critical items, and each contained two animate and two inanimate elements. Examples of visual scenes for these sentences are presented in Figure 5, and examples of filler sentences are presented below. The full set of experimental items is presented in Appendix D.
Intransitive (Unergative)

(7) H<um>ikbi marahil kanina ang anak sa guro sa silid-aralan dahlil sa pagsusulit.
    <INT.PRF>sob perhaps earlier PVT child OBL teacher OBL classroom because of exam
    ‘Perhaps the child sobbed earlier to the teacher in the classroom because of the exam.’

Intransitive (Unaccusative)

(8) Na-buhay marahil kahapon ang halaman dahlil sa insecticide mula sa sayantipiko
    INT.PRF-life perhaps yesterday PVT plant because of insecticide from scientist
    ayon sa magsasaka.
    according to farmer
    ‘Yesterday the plant perhaps became alive because of the insecticide from the scientist
    according to the farmer.’

Locative Voice

(9) T<in>ambay-an marahil kahapon ng lalaki ang upuan sa simbahan ayon sa pari.
    <PRF>-linger-LV perhaps yesterday NPVT man PVT chair OBL church according to priest
    ‘Yesterday the man perhaps lingered on the bench in the church according to the priest.’

Transitives with Experiencer Pivots

(10) Na-kinig yata kahapon ang babae ng awit sa radyo ayon sa tatay.
    PCV.PRF-listen perhaps yesterday PVT girl NPVT song OBL radio according to father
    ‘Yesterday the girl perhaps listened to the song on the radio according to the father.’

Figure 5. Visual scenes for example filler sentences 7 (top-left), 8 (top-right),
9 (bottom-left), and 10 (bottom-right).

Overall, each experimental list was comprised of 36 critical items and 72 filler items, which
totaled to 108 items. Items were pseudo-randomized, such that each critical item was separated
from the previous one by one to three filler items. In every list, no two subsequent critical items belonged to the same condition. Participants were assigned in a rotating fashion to one of these lists.

Each experimental list also began with four practice items which had a sentence structure similar to the filler items. All sentences used in the experiments were audio-recorded by a female native Tagalog speaker who was naïve to the study objectives.

2.2.2.3. Procedure

The VWP experiment utilized an SMI RED250 eye-tracker sampling at 60Hz.\(^\text{14}\) Participants sat on a height-adjustable chair while facing the eyetracker sensors and computer screen that presented the visual stimuli.

Each participant began with a nine-point calibration. All participants had calibrations of 1.5 degrees or below of the visual angle. The height of the chair and the distance of the participant was adjusted to improve the calibration process.\(^\text{15}\) During the experiment, participants were reminded from time to time to maintain the appropriate posture that was set up in the calibration process.

Each trial began with an initial preview of the visual scene for 1500 ms. An audio-recorded stimulus of the sentence items was played for approximately 1500-7000 ms, followed by a silence of another 1500 ms. A post-stimulus task followed, in which participants responded to a yes/no question about the main assertion of the sentence. An example post-stimulus question for sentence (1) was “Nagluto ba ang nanay ng spaghetti? (Did the mother cook spaghetti?)”. The syntactic structure of the questions was controlled such that both the sentence stimulus and the question were in the same voice. Participants responded to these yes/no questions using keyboard presses.

2.2.2.4. Analyses

Two participants were excluded for not obtaining an accuracy score of at least 80% on the post-stimulus task, while one participant was excluded for obtaining an extremely low score in one

---

\(^\text{14}\)Due to technical difficulties caused by a program update, the experiment was implemented under a sampling rate of 60Hz.

\(^\text{15}\)A smaller number of participants participated in the experiment with certain laboratory modifications: participants sat in a chair of a fixed height, and the eyetracker sensors and computer screen were positioned on a height-adjustable table. For these participants, the positioning of the table or the distance from the table was adjusted to attain the appropriate calibration.
of the voice conditions. Comprehension accuracy in the different voice conditions was analyzed from the remaining 46 participants.

2.2.2.4.1. Critical Items

The data were binned to 20 ms time windows and were adjusted according to the acoustic stimuli. The eye fixation data from these 46 participants were further trimmed in terms of the amount of track losses that occurred within each trial. Track losses are defined in this study as the absence of recorded looks to any of the areas of interest (AOIs) or scene background space within the spoken region (1500-7000 ms). Four participants were excluded as their average percentages of track loss across the critical trials was greater than 50%. Forty-two trials were also removed due to greater than 50% track loss in the spoken region.

Data were time-aligned such that 0 ms was set as the onset of the verb region. Proportions of looks to one of the arguments over the total number of looks to all areas of interest were computed for the critical verb + adverb region (roughly 0 to 1900 ms) and are plotted in Figure 6. The critical region was further expanded by 200 ms to account for the amount of time required to launch saccadic eye movements (Matin et al., 1993).

All binary data were converted to empirical logits by 100 ms time windows, aggregated by-participant and by-item. While such aggregation eliminates the ability to include crossed random effects for participants and items in statistical models, the use of empirical logit analyses optimizes the current data distributions that I obtained. The empirical logit values were analyzed using mixed-effects linear regression models with Voice as a fixed effect, and participants and items as random effects (Barr, 2008) using the lme4 package in R (Bates et al., 2015; R Core Team, 2013). All analyses began with the full model justified by the design, with Voice as a fixed effect, and intercepts and slopes of the fixed effects in the random effects structure. However, due to non-convergence, the random slopes were taken out from the models.

To test whether there were increased looks to the pivot element as a function of Voice during the Verb + Adverb region (Hypothesis 1), the empirical logit of looks to the pivot AOIs were used as the dependent measure. Voice was contrast-coded using treatment coding, with the argument in the pivot condition set as the reference level (e.g., looks to agent in AV; looks to patient in PV; and so on). The pair of these models that sets agent looks in AV as the reference level also
examined whether there were increased looks to the agent argument in AV, compared to the other Voice conditions (second half of Hypothesis 3). In these models, the empirical logits were computed using the following formulae, with $\epsilon$ set at 0.5:

$$\text{empirical logits} = \log\left(\frac{\text{sum of looks to X} + \epsilon}{\text{average looks to non-X entities} + \epsilon}\right)$$

$$\text{weights} = \frac{1}{(\text{sum of looks to X} + \epsilon)} + \frac{1}{(\text{average looks to non-X entities} + \epsilon)}$$

To examine whether there were overall increased looks to the agent across all conditions (Hypothesis 2), the same models specified above were constructed, with the exception that simple coding was used. Simple coding permitted the testing of the grand mean intercept, i.e., the overall looks to agents. Further, models were also planned to test for greater looks to agent versus the pivot across all conditions (first half of Hypothesis 3), by setting looks to agent in non-agent voices as the reference level. However, visual inspection of the results showed a significant animacy effect; thus, these models were not implemented.

2.2.2.4.2. Filler Items

A similar analysis was also employed for the fillers for further data analysis. They were also binned to 20 ms time windows and adjusted according to the acoustic stimuli. Further trimming of the eye fixation data from these 46 participants resulted in exclusion of 5 participants as their average percentage of track loss across the critical trials was greater than 50%. An additional 174 trials were also removed due to greater than 50% of track losses in the spoken region.

Because, as detailed below, the critical items revealed no effect of voice but a substantial effect of animacy, further analyses were employed with the fillers, with the aim of determining whether they too would exhibit a significant bias for looks to the animate entities. Thus, proportions of looks to the animates versus inanimates were computed. Data were time-aligned such that 0 ms was set as the onset of the verb region, and the entire trial duration was set as the region of interest.

The binary data from the fillers were converted to empirical logits by 100 ms time windows, aggregated by-participant. The empirical logit values were analyzed using mixed-effects linear regression models with Item Type (critical/filler) as a fixed effect, and participant as a random effect (Barr, 2008) using the lme4 package in R. Due to non-convergence, the random slopes were
taken out from the models. A linear regression model was also built with the aggregation of empirical logits by-item, with Item Type (critical/filler) as a predictor. No random effects structure was built for this model as the data structure does not allow for Item as a random effect. Item Type was contrast-coded using deviation coding to determine whether there was a significant difference between critical and filler items in the bias to look at the animate entities over the inanimate ones.

2.2.3. Results
2.2.3.1. Comprehension Accuracy

An analysis of the accuracy of responses in the post-stimulus task revealed that participants committed more errors in the benefactive and instrumental voice, compared to the agent and patient voice conditions (Table 6). Similar to the findings in Experiment 1, the higher error rates in the benefactive and instrumental voice might be attributed to their relatively lower frequency, higher processing difficulty, or need for a more constrained discourse context.

Table 6. Percentage of correct responses in the post-stimulus task across conditions (numbers in parentheses indicate standard deviation values in percentages).

<table>
<thead>
<tr>
<th></th>
<th>AV</th>
<th>PV</th>
<th>BV</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>% correct</td>
<td>96.14%</td>
<td>96.62%</td>
<td>88.89%</td>
<td>89.86%</td>
</tr>
<tr>
<td></td>
<td>(19.30%)</td>
<td>(18.10%)</td>
<td>(31.46%)</td>
<td>(30.23%)</td>
</tr>
</tbody>
</table>

2.2.3.2. Eye Gaze Data
To assess the fixations to the targeted elements relative to voice, the results were plotted as proportions of fixations to the different arguments as the sentence unfolded (Figure 6).
Figure 6. Proportion of looks to AOIs depicting the agent, benefactive, instrument, and patient arguments, by voice condition. Time is aligned with the onset of the verb. The solid box indicates the verb-adverb region. Subsequent argument regions are marked by dotted vertical lines; the pivot label is fully majuscule.

Prior to speech onset (-1500 to 0 ms), participants were observed to have a high proportion of looks to the animates (agents and benefactives) compared to the inanimates (patients and instruments) in all conditions. This bias for looks to the animates was maintained during the verb + adverb region (0-2100ms), and extends throughout the trial in all conditions.

Table 7 summarizes the results of the mixed-effects linear regression models used in the verb + adverb region, following the aggregations by-participant and by-item. To evaluate whether there were increased looks to the pivot element in the verb + adverb region, the treatment coding for each model allowed comparison of whether there were greater looks to an element when it was pivot-marked than when it was not (see Hypothesis 1). All comparisons showed non-significant differences between conditions, suggesting that increased looks to the pivot were not observed in the experiment, contra the prediction of Hypothesis 1. Nor was there a strong pattern of looks to elements as they were mentioned. Instead, a strong animacy effect was observed throughout the
spoken trial in all conditions (Figure 6). The positive beta values for the intercepts of animate AOIs (agents and benefactives) versus the inanimate AOIs (patients and instruments) suggests the dominating proportion of looks to animates than the inanimates during the verb + adverb region. Thus, despite the significant looks to the agent during the verb region, there were also significant looks to the benefactive, which contradicts the prediction in Hypothesis 2.

Table 7. Results of mixed effects linear regression models that tested for a preference to look at the pivot-associated AOI during the critical verb + adverb region, for each AOI.16

<table>
<thead>
<tr>
<th></th>
<th>By-participant aggregation</th>
<th></th>
<th>By-item aggregation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>β</strong></td>
<td><strong>SE</strong></td>
<td><strong>t</strong></td>
<td><strong>p</strong></td>
</tr>
<tr>
<td><strong>Looks to the Agent (with AV as a baseline)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>0.57</strong></td>
<td><strong>0.09</strong></td>
<td><strong>6.42</strong></td>
<td>&lt; .001</td>
</tr>
<tr>
<td>AV vs PV</td>
<td>0.06</td>
<td>0.11</td>
<td>0.55</td>
<td>0.58</td>
</tr>
<tr>
<td>AV vs BV</td>
<td>0.06</td>
<td>0.11</td>
<td>0.59</td>
<td>0.80</td>
</tr>
<tr>
<td>AV vs IV</td>
<td>0.03</td>
<td>0.11</td>
<td>0.30</td>
<td>0.80</td>
</tr>
<tr>
<td><strong>Looks to the Patient (with PV as a baseline)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>-0.80</strong></td>
<td><strong>0.14</strong></td>
<td><strong>-5.87</strong></td>
<td>&lt; .001</td>
</tr>
<tr>
<td>PV vs AV</td>
<td>-0.18</td>
<td>0.13</td>
<td>-1.35</td>
<td>0.18</td>
</tr>
<tr>
<td>PV vs BV</td>
<td>-0.07</td>
<td>0.13</td>
<td>-0.54</td>
<td>0.59</td>
</tr>
<tr>
<td>PV vs IV</td>
<td>-0.21</td>
<td>0.13</td>
<td>-1.56</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Looks to the Benefactive (with BV as a baseline)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>0.63</strong></td>
<td><strong>0.09</strong></td>
<td><strong>6.96</strong></td>
<td>&lt; .001</td>
</tr>
<tr>
<td>BV vs AV</td>
<td>0.03</td>
<td>0.11</td>
<td>0.27</td>
<td>0.79</td>
</tr>
<tr>
<td>BV vs PV</td>
<td>-0.02</td>
<td>0.11</td>
<td>-0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>BV vs IV</td>
<td>0.06</td>
<td>0.11</td>
<td>0.54</td>
<td>0.59</td>
</tr>
<tr>
<td><strong>Looks to the Instrument (with IV as a baseline)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>-1.07</strong></td>
<td><strong>0.14</strong></td>
<td><strong>-7.37</strong></td>
<td>&lt; .001</td>
</tr>
<tr>
<td>IV vs AV</td>
<td>0.20</td>
<td>0.13</td>
<td>1.46</td>
<td>0.17</td>
</tr>
<tr>
<td>IV vs PV</td>
<td>0.10</td>
<td>0.14</td>
<td>0.71</td>
<td>0.48</td>
</tr>
<tr>
<td>IV vs BV</td>
<td>0.001</td>
<td>0.14</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

16Mixed effects linear regression models used

```r
lmer(elogit_AgenttoOthers ~ Voice + (1|Participant), weights=1/wtsAgenttoOthers)
lmer(elogit_AgenttoOthers ~ Voice + (1|Item), weights=1/wtsAgenttoOthers)
lmer(elogit_PatienttoOthers ~ Voice + (1|Participant), weights=1/wtsPatienttoOthers)
lmer(elogit_PatienttoOthers ~ Voice + (1|Item), weights=1/wtsPatienttoOthers)
lmer(elogit_BenefactivetoOthers ~ Voice + (1|Participant), weights=1/wtsBenefactivetoOthers)
lmer(elogit_BenefactivetoOthers ~ Voice + (1|Item), weights=1/wtsBenefactivetoOthers)
lmer(elogit_InstrumenttoOthers ~ Voice + (1|Participant), weights=1/wtsInstrumenttoOthers)
lmer(elogit_InstrumenttoOthers ~ Voice + (1|Item), weights=1/wtsInstrumenttoOthers)
```
2.2.3.3. Post-hoc Analysis: Critical items versus Fillers
Because there was a surprisingly strong animacy effect in the critical items, I examined whether the same animacy effect was found in the filler items. I examined (i) the proportion of looks to the animate versus the inanimate entities throughout the sentence trial for the criticals versus the fillers; and (ii) the looks to an element as it was mentioned in a subset of the critical and filler items.

In contrast to the critical items, the fillers did not show a trial-wide bias for looks to the animate arguments. This is shown in Figure 7, where looks to the inanimates dominated the gaze patterns at some part of the sentence in each the different filler sentence types. The finding of a different pattern in the filler items than in the critical items is supported by the results of the linear regression models (both by-participant and by-item) which showed a significant effect of Item Type, suggesting that there was a difference with regard to the proportion of looks to the animates versus the inanimates between the critical and filler items (Table 8).

Figure 7. Proportion of looks to animate/inanimate entities in the filler items, by sentence type. Time is aligned with the onset of the verb.
Table 8. Results of the linear regression models that tested the proportion of looks to animates versus inanimates.\textsuperscript{17}

<table>
<thead>
<tr>
<th></th>
<th>By-participant aggregation (lmer)</th>
<th>By-item aggregation (lm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>SE</td>
</tr>
<tr>
<td>Looks to Animates vs Inanimates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.83</td>
<td>0.06</td>
</tr>
<tr>
<td>Item Type</td>
<td>-0.43</td>
<td>0.09</td>
</tr>
</tbody>
</table>

For the subset analysis, I examined the looks to one animate and one inanimate element as it was mentioned in the sentence to determine whether the critical and filler items differed in the effect of animacy (Figure 8). For the critical items, I chose the looks to the patient and benefactive arguments as they were mentioned in the NP2 (2800-3800ms) and NP3 (3800-4800ms) regions of the PV and BV sentences (Figure 8, left panel). In the PV condition, the patients were mentioned in NP2, while the benefactives were mentioned in NP3. In the BV condition, the mention to the patients and benefactives in NP2 and NP3 were in the reverse order. I determined whether there would be an interaction effect between Voice and Region (both sum-coded) that would capture the changes in the looks of these elements as they are mentioned (Table 9). There was a significant interaction effect between Voice and Region with regard to looks to the patient in both by-participant ($p < .01$) and by-item aggregations ($p < .05$), suggesting that there were increased looks to the inanimate patient as it was mentioned. However, there was not a significant interaction effect for looks to the animate benefactive; it remained non-significant in the by-participant ($p = 0.08$) and the by-item aggregations ($p = 0.23$), suggesting that there were no major changes to the looks to the animate benefactive based on mention. Thus, for the critical items, while there were some changes to the looks to the inanimate entities, the strong bias for looks to the animate elements was maintained as the sentences unfolded.

\textsuperscript{17}Mixed effects linear regression models used

\begin{verbatim}
lmer(elogit_AnimatetoOthers ~ ItemType + (1|Participant), weights=1/wtsAnimatetoOthers)
lm(elogit_AnimatetoOthers ~ ItemType, weights=1/wtsAnimatetoOthers)
\end{verbatim}
Figure 8. Subset analysis involving the looks to mentioned/unmentioned elements in selected sentence regions for the critical items (left) and the filler items (right). For the critical items, the looks to the patients and benefactives are examined in NP2 and NP3 of PV and BV sentences as they are mentioned. For the filler items, the looks to comparable mentioned entities, coded as Animate1 and Inanimate4, are examined in NP2 and NP3 on a subset of unaccusative sentences.

Table 9. Results of the linear regression models that examined the looks to patients/benefactives as they were mentioned in NP2/NP3 in PV and BV sentences.\(^\text{18}\)

<table>
<thead>
<tr>
<th></th>
<th>By-participant aggregation (lmer)</th>
<th>By-item aggregation (lmer)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\beta)</td>
<td>SE</td>
</tr>
<tr>
<td><strong>Looks to the Patient</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>-0.46</strong></td>
<td>0.09</td>
</tr>
<tr>
<td>Voice</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Region</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td>Voice*Region</td>
<td><strong>-0.55</strong></td>
<td>0.20</td>
</tr>
<tr>
<td><strong>Looks to the Benefactive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>0.53</strong></td>
<td>0.07</td>
</tr>
<tr>
<td>Voice</td>
<td>-0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>Region</td>
<td>0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>Voice*Region</td>
<td>0.29</td>
<td>0.17</td>
</tr>
</tbody>
</table>

\(^{18}\)Mixed effects linear regression models used (Voice and Region contrast coded)

- `lmer(elogit_PatienttoOthers ~ Voice*Region + (1+Voice+Region|Participant), weights=1/wtsPatienttoOthers)
- `lmer(elogit_PatienttoOthers ~ Voice*Region + (1+Voice+Region|Item), weights=1/wtsPatienttoOthers)
- `lmer(elogit_BenefactivetoOthers ~ Voice*Region + (1+Voice+Region|Participant), weights=1/wtsBenefactivetoOthers)
- `lmer(elogit_BenefactivetoOthers ~ Voice*Region + (1+Voice+Region|Item), weights=1/wtsBenefactivetoOthers)`
For the filler items, I examined the looks to the one of the animates (labelled “animate1”) and inanimates (“inanimate4”) as they were mentioned in the NP2 (3100-4000ms) and NP3 (4000-5400ms) regions of a subset of the unaccusative sentences (Figure 8, right panel). In the first word order type, animate1 was mentioned in NP2, while inanimate4 was mentioned in NP3. These elements were mentioned in the reverse order for the second word order type. Similar to the previous subset analyses, I determined whether there would be an interaction effect between Voice and Region (both deviation-coded) that would capture the changes in the looks of these elements as they are mentioned (Table 10). There was a significant interaction effect between Voice and Region with regard to looks to animate1 ($p < .001$) and inanimate4 ($p < .001$) in the by-participant aggregation, suggesting that there was an increase in the looks to both the animates and the inanimates as they were mentioned in the sentence. Neither interaction effect reached significance in the by-item aggregation, which I attribute to the small number of included items – just 3 for each filler word order in the subset analyses. Nevertheless, the filler results overall suggest that looks to both the animates and inanimates in the filler items were sensitive to mention, and that the animacy preference was stronger in the critical items than the filler items. They further suggest that the strong animacy preference in the critical items was not simply due to an experiment-wide bias, and that gaze patterns in the experiment as a whole were responsive to the spoken material. I discuss the unexpected gaze patterns of the critical items further in the subsequent section.

Table 10. Results of the linear regression models that examined the looks to Animate1/Inanimate4 as they are mentioned in NP2/NP3 in a subset of unaccusative sentences. 

<table>
<thead>
<tr>
<th></th>
<th>By-participant aggregation (lmer)</th>
<th>By-item aggregation (lmer)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>SE</td>
</tr>
<tr>
<td>Looks to Animate1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.24</td>
<td>0.07</td>
</tr>
<tr>
<td>WordOrder</td>
<td>-0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>Region</td>
<td>0.11</td>
<td>0.14</td>
</tr>
<tr>
<td>WO*Region</td>
<td>0.97</td>
<td>0.27</td>
</tr>
</tbody>
</table>

19 Mixed effects linear regression models used (WordOrder and Region contrast coded)

lmer(elogit_Anim1toOthers ~ WordOrder*Region + (1+WordOrder+Region|Participant), weights=1/wtsAnim1toOthers)
lmer(elogit_Anim1toOthers ~ WordOrder*Region + (1|Item), weights=1/wtsAnim1toOthers)
lmer(elogit_Inanim4toOthers ~ WordOrder*Region + (1+WordOrder|Participant), weights=1/wtsInanim4toOthers)
lmer(elogit_Inanim4toOthers ~ WordOrder*Region + (1|Item), weights=1/wtsInanim4toOthers)
2.2.4. Discussion

Experiment 2 examined the influence of agent and pivot prominence on comprehension by determining whether comprehenders use voice information on the verb to anticipate the two most likely upcoming arguments, i.e., the agent and the pivot, upon hearing the verb, and preferentially fixate on the images representing these arguments in the visual scene. The results of mixed-effects linear regression models on fixation patterns revealed no significant effect of Voice on looks to agent or pivot AOIs in the critical verb + adverb region for any of the voices. In contrast to the predictions, the results failed to demonstrate the presence of a Voice effect in anticipating the likely arguments upon hearing the verb. The lack of preferential looks to the pivot as comprehenders received the verb and voice information was surprising, considering the strong effect of pivot prominence on linear word order that was found in production, the tight linguistic relationship between voice-marking and the pivot, and the well-documented role of syntactic prominence in theoretical linguistic work.

In addition, despite extensive cognitive and linguistic processing research that has established agent saliency, there was a lack of preferential looks to the agent upon hearing the verb and voice morphology. There were many looks to the agent in the verb + adverb time window, yet the proportions of looks to the agent were comparable to those for the benefactive argument, the only other animate entity in the display. Given the strong looks to both animate entities, it is difficult to postulate that there were distinctive preferential looks to agents as a function of verb and voice.

Furthermore, there was a general lack of differentiation across the four voice conditions in looks during the unfolding of the first two arguments. Experiment 1 illustrated that non-agent voices are likely to have the agent and pivot as their first two arguments respectively, in contrast to the agent voice, where the agent pivot could either be in the first or second position, thus producing an Agent-Patient or Patient-Agent word order. If gaze patterns in comprehension had mirrored the word order patterns in production, the result would have been (i) a split across

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.34</td>
<td>0.07</td>
<td>4.92</td>
<td>&lt;.001</td>
<td>0.35</td>
<td>0.19</td>
<td>1.83</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>WordOrder</td>
<td>-0.15</td>
<td>0.14</td>
<td>-1.06</td>
<td>0.29</td>
<td>-0.20</td>
<td>0.38</td>
<td>-0.51</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>-0.26</td>
<td>0.14</td>
<td>-1.86</td>
<td>0.06</td>
<td>-0.23</td>
<td>0.38</td>
<td>-0.61</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>WO*Region</td>
<td>-0.98</td>
<td>0.28</td>
<td>-3.52</td>
<td>&lt;.001</td>
<td>-1.27</td>
<td>0.76</td>
<td>-1.66</td>
<td>0.15</td>
<td></td>
</tr>
</tbody>
</table>
conditions in looks to the first argument, between a preference the agent for the non-agent voice patterns, versus either the agent or the patient for the agent voice; and (ii) a strong preference for looks to the pivot in conjunction with the second argument, especially for the non-agent voice patterns, differentiating the patient, benefactive, and instrument voices. Additional models (not discussed for reasons of space) established that there was no split in the looks to the first argument, before or during its mention, nor were there preferential looks to the pivot at the second argument. The gaze patterns for all voice patterns are very similar.

One might wonder if an effect of agent and pivot prominence was not observed as comprehenders received the verb and voice information because of the time needed to process the verbal information. The parser might require considerable time to process the rich verbal information sufficiently to produce predictions of high granularity. Such time-sensitivity has been demonstrated in previous research, where successful predictive effects were observed only when the parser was provided with more time to process and predict the input (Chow et al., 2018; Chow, Smith, et al., 2016; Huettig & Guerra, 2019; Ito et al., 2016). However, the experimental stimuli tested here provided the comprehenders with ample time to process the information; the verb and adverb region lasted for about 1900ms, with an additional extension of 200ms in the analysis for saccadic eye movement. Comprehenders had at least two seconds to process the information and launch eye movements that could be reflective of their likely predictions.

A more convincing explanation for the lack of significant effects of agent and pivot prominence may be related to the quantity of information needed for the parser to generate predictions about the upcoming signal. In this study, I tested predictive effects of agency and pivothood when comprehenders had only been provided with a verb and (unconstraining) adverb. Research on predictive processing in the literature is dominated by experimental work that offers long and informative sentential contexts (e.g., DeLong et al., 2005; Federmeier & Kutas, 1999; Huettig & Guerra, 2019), or the use of highly predictable patterns such as either... or... (e.g., Staub & Clifton, 2006). In addition, previous studies that examined effects of using verb information to anticipate likely grammatical objects (e.g., Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003), or of using verb subcategorization information to predict the upcoming structure (e.g., Arai & Keller, 2013; Boland et al., 1995; Linzen & Jaeger, 2016; Trueswell et al., 1993) had at least one argument (such as the sentential subject) along with the verb to facilitate the generation
of predictions; it may be the case that more information, such as an argument in addition to the verb and voice information, would have been necessary for comprehenders to make predictions about the upcoming signal in the present study. This point is discussed further in the General Discussion chapter.

Instead of finding an effect of agency and pivoting, what was consistently observed in the gaze patterns in all conditions was a strong animacy effect that lingered throughout the spoken trial. The strong influence of animacy with anticipatory gaze patterns was unexpected, as animacy did not have a critical role in linear word order in Experiment 1.

Could the strong animacy effect in the verb + adverb region actually be an effect of agent prominence, in the sense that comprehenders were possibly considering each of the animates as a potential agent? Such an interpretation could converge with the outcomes of Sauppe (2016), who reported greater looks to the agent element upon hearing the voice-marked verb, regardless of the voice condition. He interpreted the results of his comprehension eyetracking experiment as showing that participants were using verbal semantics over morphosyntactic information to anticipate the upcoming input upon hearing the verb. He also interpreted the overall greater looks to the agent element in both voice conditions as a sign of agent prominence. However, an agent prominence interpretation is undercut by the norming results, in which the intended agent in the visual scene was identified by the participants with at least 80% accuracy in each voice condition. Coupled with the ample time given to participants to inspect the scene prior to the onset of the sentence, it seems unlikely that they struggled with agent identification. In addition, the agent prominence explanation does not explain why comprehenders maintained strong looks to both animate arguments throughout the sentence, including after the agent was mentioned in the NP1 region. Even in the agent voice condition, in which morphosyntactic information marks the relevant argument as the agent, high proportions of look to the two animate arguments persisted to the end of the sentence. Hence, I deem it unlikely that the frequent looks to both animate entities could have been driven solely by agent prominence. More likely, in my view, is that a strong animacy effect influenced anticipatory eye gaze patterns, in the present results and perhaps in Sauppe’s as well.

Recall that the visual scenes counterbalanced the left-right position of the arguments, but consistently placed the animate entities in the top row and the inanimate ones in the bottom row.
One could ask whether the dominant gaze patterns in Experiment 2 were attributable to lower-level attentional effects for the top row, to visual salience for animate entities, or to linguistic processing of animates as more prominent arguments. To investigate this question, I conducted further analysis to examine whether the strong animacy effect would also be found in the filler items, which employed similar scenes to the critical items, with two animate entities in the top row and two inanimate ones in the bottom row. An examination of the proportion of looks to animates versus inanimates between critical and filler items showed a clear asymmetry. At appropriate points in the sentence for linguistically driven looks, comprehenders had greater looks to the inanimates than the animates in the filler items, unlike in the critical items where the eye gaze patterns were dominated by looks to the animates throughout the trial. The disparity between the criticals and fillers was further demonstrated by the subset analysis that looked at gaze patterns of mentioned/unmentioned elements in the NP2/NP3 regions of selected sentence types in the critical and filler items. This asymmetry casts doubt on attention or visual saliency fully accounting for the patterns, and supports the conclusion that linguistic processing, along with an animacy preference, influenced the eye gaze patterns in the critical items.

If linguistic processing shaped gaze patterns, why did the critical items produce largely undifferentiated patterns across voice conditions? I speculate that it was due to the wide range of predictions that the comprehenders were making. All of the entities depicted in the visual scenes of the critical items could be predicted from the verb. All four entities in these visual scenes were both semantically related to the verb (e.g., in Figure 1, the agent, patient, benefactive, and instrument are all relevant to the event of cooking), as well as syntactically related to it (i.e., all of the entities in the visual scene could be treated as the sentential pivot using one voice alternation or another), and all could occur as arguments within a single clause. This was not the case for the filler items. In filler sentences, only one or two entities were predictable from the verb, and the other entities were unrelated (e.g., source, cause, or accordant). Thus, even though comprehenders may be predicting the agent and the pivot in the first and second argument positions of the critical items, they could also be generating additional predictions for the sentence that is likely to unfold. With this abundance of potential predictions, gaze patterns might not converge reliably on the agent or the pivot. The outcomes observed in this experiment contrast with the abundant literature involving visual world eyetracking that has examined the role of prediction in the comprehension
of languages with restricted word orders, where the initial subject/agent determines initial looks, and the critical or target element occurs after more linguistic information had accumulated.

The variability of predictions that I postulate for Experiment 2 may wash out distinct gaze patterns for each Voice condition throughout the trial. Although comprehenders could have potentially generated sharper predictions about later parts of the unfolding signal (e.g., elements about to be mentioned) as they integrated information from earlier parts of the input (i.e., elements mentioned from the signal), they were afforded only limited time in the nominal regions to launch saccadic eye movements that would have reflected predictions to the rest of the unfolding sentence. Since the focus was on predictions made in response to voice marking on the verb, the experimental stimuli (the nominal arguments in particular) were not designed for the examination of anticipation to the unfolding nominal arguments in these later regions.

The amount of linguistic experience comprehenders accrue may also contribute to the variability of predictions: greater experience with a particular pattern could boost the strength of predictions for that pattern. Some support for this comes from the subset analysis that examined looks to entities as they were mentioned (for two conditions that could sensibly be compared). In the critical items, I only found an effect of the looks to the inanimate patient element in the subset analysis. Experiment 1 showed that the highly frequent patient voice had the greatest consistency in word order, with a strong preference for the patient pivot in the second position. In the subset analysis, the patient voice garnered increased looks in the second argument position and decreased looks in the third argument position, congruent with the idea that stronger predictions drive more looks to entities that are about to be mentioned. In contrast, the production patterns, and so presumably the predictions, for the animate benefactive pivot in the less common benefactive voice were less strong compared to the patient voice; correspondingly, the looks to the animate benefactive element in the subset analysis did not change significantly in response to its mention in the sentence.

The findings from Experiment 2 call for further exploration of the link between behavioral measures such as gaze patterns and anticipation in comprehension. The results of the current experiment highlight effects on prediction of the number of sentential arguments in the unfolding signal, the use of longer, more restrictive sentential contexts, and the amount or uniformity of
experience with a construction. Future research could also focus on distinguishing the effect of agent prominence from animacy by manipulating the animacy of sentential arguments.

2.3. General Discussion

This study examined the influences of agent prominence, syntactic prominence or pivothood, and animacy, in linear word order in production and in anticipatory gaze patterns in comprehension. These constraints were examined in a verb-initial language, Tagalog, that has a voice system that allows these factors to be teased apart. Experiment 1 tested two constraints rooted on agent and pivot prominence (namely agent-first and pivot-second constraints), which were postulated to drive word order preferences in the language. The outcomes showed that speakers produce patterns that align well with the hypothesized agent-first and pivot-second constraints, realized in the agent voice such that the agent can be in either the first or second position. This result confirmed the influence of agent and pivot prominence in production, with both of these factors bearing comparable strength on linear word order. The influence of animacy in linear word order was minimal, as animacy did not affect the strong preference for speakers to position the pivot in the second argument position.

I then determined whether agent and pivot prominence would also impact real-time anticipation of two likely sentential arguments of the verb, i.e., the agent and the pivot, as comprehenders received verb plus voice information in the linguistic signal. The exact opposite of the constraint influences in Experiment 1 was observed in Experiment 2. Influences of agent and pivot prominence were not demonstrated in the comprehenders’ anticipatory gaze patterns. Instead, a strong animacy effect was observed throughout the critical trials. I attribute the lack of an anticipatory effect of agency and pivothood to the limits of predictive processing: a sufficient quantity of information may be necessary for the parser to successfully generate strong and uniform predictions about the upcoming signal. These findings distinguish the current study from much past work on predictive processing, which has provided rich sentential contexts that permit successful prediction. The current findings indicate some of the limits on predictive processing as we critically examine its role in information processing.

I believe that the findings in Experiment 2 do not downweigh the potential effects of agent and pivot prominence in sentence processing; the influences of agent and pivot prominence have been
solidly documented in the growing research on Tagalog. For example, agent prominence has been
observed in work by Bondoc et al. (2018) and N. Tanaka et al. (2019) on relative clauses, where
Tagalog speakers of different populations showed an advantage for relative clauses that involve
an agent as the head. There is also evidence for processing advantages in declarative sentences that
involve the agent-first word order, regardless of voice, that points towards agent prominence
(Bondoc et al., 2018; Garcia et al., 2020; Garcia, Roeser, et al., 2018). The use of voice
morphology in processing filler-gap dependencies has also been demonstrated in a recent study by
Pizarro-Guevara and Wagers (2020). They showed that Tagalog comprehenders demonstrated
predictive effects of identifying the gap site after encountering the head of a relative clause and
the verb containing the voice morphology, although the strength of predictive effects varied
depending on the type of filler-gap dependency. These past studies, however, did not look at real-
time comprehension effects when comprehenders were only afforded the verb and voice
morphology prior to making a prediction. I emphasize the need to investigate in future research
whether predictive effects vary depending on the type of syntactic phenomena, as well as the
sources of linguistic information available in the signal.

An important finding in the study is the difference with regard to constraint influences between
the two experiments. To recap, in Experiment 1, agent and pivot prominence, but not animacy,
had major influences on syntactic linear word ordering in production. In Experiment 2, I found an
animacy effect, but not agent or pivot prominence effects, on gaze patterns in comprehension. A
similar production-comprehension mismatch with regard to agent and pivot prominence was found
in the results reported by Sauppe (2016) and Sauppe and colleagues (2013) on Tagalog, which
reported a pivothood or voice effect in production but not in comprehension. These findings
indicate a potential difference between the production and comprehension systems, which runs
contrary to several proposals in the literature that unify the production and comprehension systems
(Dell & Chang, 2013; MacDonald, 2013; Momma & Phillips, 2018; Pickering & Gambi, 2018;
Pickering & Garrod, 2007, 2013). Because the speakers and the comprehenders were given the
same visual and linguistic stimuli in the present study (i.e., a verb with voice morphology, followed
by an adverb), the mismatch in the behavioral results potentially lies with the different processing
goals of the speaker and the comprehender. In production, speakers need to determine the
hierarchical relationships between the event and the lemmas as they formulate a linguistic
message. Despite the relevance of all entities in the visual scene to the event, they need to commit to a single string as they formulate and linearize the message. In comprehension, while the comprehender can make predictions on how the signal will unfold, multiple parallel probabilistic predictions could be generated about the unfolding signal. The degree of commitment to a single prediction may vary, depending on the likelihood of that prediction in a particular context. In the case of Experiment 2, I argued that the variability of the predictions affected the commitment to a single prediction, and resulted in a uniform gaze pattern across conditions, given the relevance of all entities in the visual scene. Prediction strength may also depend on the quantity of information provided in the signal. I argued that the verb information alone might not be adequate to make fine-grained predictions that can be observed in real-time. Experimental outcomes such as these that demonstrate contrasts between production and comprehension highlight the need to examine the relationship between these two systems and determine which components between the two are similar or different.

Finally, a mismatch in the use of these constraints may also reflect a possible typological difference between languages with more or less word order variability. Despite the strengths of the proposed agent-first and pivot-second constraints on word order in production, the flexibility of word order patterns in Tagalog may be a driving factor that contributes to the lack of significant anticipatory gaze patterns in comprehension. There may be a processing advantage in a free word order language for the comprehender to anticipate any argument that is semantically and syntactically relevant to the verb, resulting in the undifferentiated looks to entities in the unfolding sentence across voice conditions in Experiment 2. Further research should examine whether similar results would be found in languages with restricted word order patterns, provided with visual scenes that are similar to the ones utilized in the present experiments.

2.4. Conclusion

This chapter examined the influences of agent and pivot prominence with regard to syntactic linear word order patterns in production and anticipatory gaze patterns in comprehension. I also assessed the overall strength of pivot prominence as a constraint by manipulating the animacy features of the pivot element. The experiments demonstrated a robust influence of agent and pivot prominence in linear word order in production, but did not find corresponding patterns in
comprehension. Our findings contribute to the growing evidence for the role of these types of prominence constraints in linguistic processing, and emphasise how constraint effects could differ in production and comprehension. The next chapter investigates how voice morphology is used by comprehenders to predict the linear word order of the unfolding sentence using reading tasks.
Chapter 3

Voice Effects on the Prediction of Word Order Patterns

The previous chapter described the role of the agent and pivot prominence in linear word order patterns in Tagalog, established the link between voice morphology and word order, and examined whether comprehenders anticipate the agent and the pivot as they hear the verb with the voice morphology. This chapter addresses whether voice morphology can be used by comprehenders to predict the specific linear word order pattern of the sentence. This chapter also examines the claim regarding the pervasiveness of predictive processes in the processing system. Three reading experiments are presented in this chapter to examine these specific questions.

3.1. Experiment 3: Self-Paced Reading

3.1.1. Introduction and Hypotheses

The visual world eyetracking experiment failed to demonstrate the anticipation of the agent and pivot as a function of the verb and the voice morphology. Instead, a strong animacy effect was observed. As was discussed in Chapter 2, this result was potentially brought about by the variability of the predictions that the comprehenders were making.

While the effects of using voice morphology to anticipate the agent and the pivot were not observed in the previous experiment, a different methodology may shed light on the comprehenders’ use of verb and voice information in predicting the upcoming structure. The visual world eyetracking experiment constrained the possibilities of the likely elements in the unfolding signal, therefore the experimental set-up might have afforded the comprehenders the alternative to not prioritize the prediction of the upcoming argument or structure. A different methodology that does not restrict the alternatives of the upcoming structure (such as a reading task) may increase the probability that comprehenders would engage in prediction for processing efficiency.

The use of voice information to predict the upcoming structure also remains unresolved. The former eyetracking experiment examined whether the influence of agentivity and pivothood would encourage comprehenders to anticipate the agent and the pivot as the likely upcoming verb arguments upon hearing the verb and the voice information. An experiment with a different
methodology could explore the comprehenders’ use of voice information to predict the specific linear word order of the upcoming structure.

The following experiment used self-paced reading to probe whether native Tagalog comprehenders use voice morphology to develop syntactic expectations for specific linear word order patterns. The generation of expectations by comprehenders can be measured using computational formalisms, i.e., Bayesian surprisal, following expectation-based parsing (Hale, 2001; Levy, 2008). Expectations that align with the predicted patterns would manifest in low surprisal values, while dashed expectations produce the opposite surprisal effect. In addition, the expectations of the comprehenders are also updated as they accrue more information from the unfolding signal.

The following hypotheses were tested:

**Hypothesis 1:** Comprehenders use the voice information on the verb to develop expectations for the particular word order of the sentence.

I predicted that comprehenders can base their probabilistic expectations for the likely word order patterns that correlate with voice morphology, as reported in Experiment 1. The findings in Experiment 1 entail that it is highly probable to expect the Agent-Patient word order in the patient voice, while there is near-chance probability to expect either the Agent-Patient or the Patient-Agent word order in the agent voice. Furthermore, the probability would be low for expecting a Patient-Agent word order in the patient voice.

Following the hypothesis, these probabilistic expectations would be demonstrated as soon as comprehenders encounter the nominal marking of the first noun phrase. Translating these probabilistic expectations in terms of surprisal, the most expected word order pattern (PV VAP) would result in the lowest surprisal values (i.e., the shortest reaction times or RTs), while the least expected word order (PV VPA) would result in the highest surprisal values (i.e., highest RTs). The two agent voice patterns would have surprisal values that lie between the two patient voice patterns. Hence, the probabilistic expectations would be demonstrated in terms of a cline of surprisal values, as follows:

\[
\begin{align*}
\text{PV VA(P)} & < & \text{AV VA(P)} & = & \text{AV VP(A)} & < & \text{PV VP(A)} \\
\text{most expected} & & \text{least surprisal} & & \text{shortest RTs} & & \text{least expected} & & \text{most surprisal} & & \text{longest RTs}
\end{align*}
\]
Thus, I predict a disordinal interaction effect between these two variables based on the cline of surprisal values, which can be observed as comprehenders encounter the nominal marking of the first noun phrase.

**Hypothesis 2**: While a gradience of surprisal values is predicted based on the first encountered NP element, no surprisal is expected as comprehenders encounter the second noun phrase. The accrual of information from the verb and first noun phrase is adequate to update the expectations of the comprehenders on the unfolding second noun phrase. Given that the second argument in each of these patterns would be the only highly expected argument (more than any other argument), I hypothesize the absence of differences between conditions, as the comprehenders encounter the nominal marking in the second noun phrase. The lack of differences will be observed as similar RTs or surprisal levels in this region. I expect no main effect of voice and word order, nor an interaction effect between these two at the nominal marking in the second noun phrase.

3.1.2. Method

Self-paced reading was utilized to examine the predictive effects of voice morphology without being influenced by visual animacy. Self-paced reading also allowed for the measurement of the development of syntactic expectations for particular word order patterns using Bayesian surprisal. To test the study hypotheses, I manipulated voice and word order as factors to obtain the following four conditions: AV VAP, AV VPA, AV VPA, and PV VPA.

In this experiment, critical sentential regions—the verb, first noun phrase, and second noun phrase—were presented in increments to measure response times in real time as each fragment was encountered (Table 11). Each sentence consisted of a warm-up adjunct (Region 1), the main verb (Region 2), the initial half of the first NP (Region 3), the latter half of the first NP (Region 4), the initial half of the second NP (Region 5), the latter half of the second NP (Region 6), and an offset adjunct (Region 7). Dividing each noun phrase into two regions permitted the testing of syntactic expectations that comprehenders develop as they encounter relevant information in the sentence.
Table 11. Stimulus items for each condition in Experiment 3.

<table>
<thead>
<tr>
<th>Voice</th>
<th>Word Order</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Region 3</th>
<th>Region 4</th>
<th>Region 5</th>
<th>Region 6</th>
<th>Region 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjunct</td>
<td>V + TimeExp</td>
<td>NM + Num</td>
<td>Adj + N</td>
<td>NM + Num</td>
<td>Adj + N</td>
<td>Adjunct</td>
<td></td>
</tr>
</tbody>
</table>

Because of the accident
Dahil sa aksidente
Dahil sa aksidente

(11) AV VAP

Because of the accident
Dahil sa aksidente
Dahil sa aksidente

(12) PV VAP

Because of the accident
Dahil sa aksidente
Dahil sa aksidente

(13) AV VPA

Because of the accident
Dahil sa aksidente
Dahil sa aksidente

(14) PV VPA

The first hypothesis (the cline of surprisal values) can be examined in Region 3. The degree of surprisal in Region 3 would be reflective of the match between syntactic expectations generated from voice morphology and the first nominal marking comprehender encounters in the input. The second hypothesis, which involved the updating of syntactic expectations (absence of difference between patterns) can be tested in Region 5. At this point, the comprehenders, who have accrued information about the verb and the first noun phrase, should have updated their expectations congruent with the highly expected subsequent argument. When they encounter the second nominal marking, the reaction times or surprisal values should not show any differences across the four sentence patterns as a result of their updated beliefs.

To maintain relatively equal lengths between regions, a numeral was added in Regions 3 and 5, and an adjective was added in Regions 4 and 6. The second half of every NP region served as a potential spillover region. Some integrative processing in the form of computing the head-dependent relations is also potentially involved in these regions, which will be discussed further in the results section.
3.1.2.1. Participants
Seventy-one native Tagalog speakers (age range 18-48, mean 25.89; 41 females, 30 males) from Metropolitan Manila in the Philippines participated in the experiment. None of the participants had any visual or hearing issues that impeded task performance. All participants identified themselves as Tagalog-dominant in a language background questionnaire. All participants also spoke English as a second language, in varying degrees of proficiency. Some participants also identified themselves as speaking a third or fourth language, but of lower proficiency compared to Tagalog and English. Regarding the Tagalog dialects spoken by participants, 37 identified as speakers of the Metro Manila dialect only, 14 as speakers of a northern Tagalog dialect only, 12 as speakers of a southern Tagalog dialect only, and eight as speakers of the Metro Manila dialect and another northern or southern Tagalog dialect. None of these participants were linguistics majors.

3.1.2.2. Materials: Development of Stimuli Items
Fifty-two verbs for the norming study were drawn from McFarland's (1989) frequency count of Tagalog lemmas. The verbs were controlled in terms of their transitivity and the elements that they take as pivots. All 52 verbs took animate agent and patient elements as pivots and were predominantly realized as transitive verbs.\(^\text{20}\)

Complete sentences containing an agent and a patient argument were created from these 52 verbs. Highly frequent nominal elements based on the McFarland frequency count were also used for the sentences. To ensure the grammatical soundness and naturalness of the sentences for the experiment, a grammaticality judgment task was implemented, in which native Tagalog participants were asked to rate the naturalness of the sentences for the experiment using a 10-point Likert scale. These 52 verbs were crossed with two voice alternations and the two word order patterns in a Latin-square design, distributed to four lists. An additional 48 filler sentences were also interspersed among these critical items. These filler sentences included 18 grammatically acceptable sentences of other sentence types and 30 grammatically unacceptable sentences

\(^\text{20}\)Because most Tagalog verbs can either be affixed by transitive or intransitive morphology, their predominant transitive use was evaluated in terms of their frequency to be used as a transitive (versus as an intransitive) in McFarland’s frequency count.
(incorrect word order patterns, incorrect voice morphology, or missing nominal marking). All sentences within each list were pseudo-randomized.

The judgment ratings were aggregated by verb, voice, and word order, such that there was a rating for each verb for each condition. These ratings were transformed to z-scores, and verbs that had z-scores lower than -0.5 in at least one condition were excluded. A total of 36 verbs were selected on the basis of the norming study.

3.1.2.3. Materials: Experimental Items

The experiment consisted of 36 critical item verbs, crossed with the two voice alternations and the two word order patterns in a Latin-square design, distributed to four lists. An additional 72 filler items were used, consisting of intransitives or other sentential patterns (e.g., benefactive/instrumental voice and complex sentences). Sentences in each list were pseudo-randomized to ensure that no two critical items of the same condition occurred consecutively.

Post-stimulus yes/no comprehension questions were also prepared to test participant comprehension. A sample post-stimulus question for sentence (5) was: “Humanap ba ang mga pamilya ng mga doktor? (Did the families look for doctors?).” The syntactic structure of the questions was controlled so that both the sentence stimulus and the question were in the same voice and word order.

In order to test whether the participants understood the main assertion of the sentence, half of the questions in the critical items (18/36 items) involved thematic role reversals by switching the position of the nouns in the sentence. One-sixth of the questions in the filler items (12/72 items) also involved thematic role reversal by switching the nominal marking of the nouns in the sentence. Finally, a third of the questions (24/72 items) involved changing the nominal or the verb with a phonological or semantic associate. The full set of experimental items is presented in Appendix E.

3.1.2.4. Procedure

The self-paced reading (non-cumulative moving window) experiment was implemented using the program E-Prime (Schneider et al., 2002) on a laptop. Participants were assigned in a rotating fashion to one of the lists. They were asked to press a key to read the next sentence fragment, in such a way that the program recorded the online response time in reading and processing each
sentence fragment. Participants responded to the post-stimulus task (yes/no comprehension question) by pressing specific letters on the keyboard.

3.1.2.5. Analyses
The comparison between the accuracy of the participants in the post-stimulus task were analyzed using a mixed effects logistic regression model in R (R Core Team, 2013) using the lme4 package (Bates et al., 2015). The analysis began with the maximal model justified by the design (Barr et al., 2013), with the binary comprehension responses as the dependent variable, with Voice and Word Order as the fixed effects, and with participants and items as random effects. Due to non-convergence, Word Order was nested within Participant, while the intercept and slope for Voice was included for Item. To test for the main effects of Voice, Word Order, and the interaction between these two variables for the relevant regions, deviation coding was used, which set the grand mean as the intercept.

For the online task, the reaction times (in ms) for each region were extracted for every condition. Critical items with incorrect responses were excluded. The reaction times with the remaining data were log-transformed, and the reaction times that had values greater than 3 standard deviations within the means, separated by region, were excluded. The log-transformed reaction time values for every sentence region were statistically treated using a maximal mixed effects linear regression model in R (Barr et al., 2013; R Core Team, 2013) using the lme4 package (Bates et al., 2015), with the Reaction Times as the dependent variable, Voice and Word Order as predictors, and with Items and Participants as random effects. Maximal models were built by including intercepts and slopes of the fixed effects in the random effects structure. Terms were removed in some of the models due to non-convergence. Similar to the analysis on accuracy responses, deviation coding was used, such that the grand mean was the intercept.

3.1.3. Results
This section presents the results from the post-stimulus task (yes/no questions) in the Self-Paced Reading Experiment (Experiment 3), followed by the region-by-region reaction time results.
3.1.3.1. Comprehension Accuracy

Three participants were excluded due to low post-stimulus task accuracy scores (<65%). Only data from 68 participants were included in the analysis (Figure 9).

![Comprehension Accuracy Between Conditions](image)

**Figure 9.** Comprehension accuracy rates in the post-stimulus task. Error bars reflect 95% CI.

In terms of voice, accuracy rates were higher in the patient voice condition ($\bar{x} = 80.56\%$, $sd = 0.40$) compared to the agent voice condition ($\bar{x} = 77.70\%$, $sd = 0.42$), and the difference between these two conditions was significant ($z = 2.62$, $p < 0.01$). As for word order, participants had better performance in the Agent-Patient order ($\bar{x} = 86.93\%$, $sd = 0.34$) compared to the Patient-Agent pattern ($\bar{x} = 71.32\%$, $sd = 0.45$), and a significant difference between these two word order patterns was observed ($z = -5.16$, $p < 0.001$). Looking at the cross between voice and word order conditions, the participants obtained the highest accuracy rate in the patient voice condition of the Verb-Agent-Patient pattern ($\bar{x} = 88.89\%$, $sd = 0.36$), followed by the agent voice condition of the same word order ($\bar{x} = 84.97\%$, $sd = 0.31$). There was no significant interaction effect between Voice and Word Order ($z = 1.28$, $p = 0.20$). Table 12 summarizes the results of the mixed-effects logistic regression model.
Table 12. Results of the mixed effects logistic regression models for comprehension accuracy.\textsuperscript{21}

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>SE</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.77</td>
<td>0.10</td>
<td>18.46</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>0.49</td>
<td>0.19</td>
<td>2.62</td>
<td>0.01</td>
</tr>
<tr>
<td>Word Order</td>
<td>-0.73</td>
<td>0.14</td>
<td>-5.16</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice:Word Order</td>
<td>0.33</td>
<td>0.25</td>
<td>1.28</td>
<td>0.20</td>
</tr>
</tbody>
</table>

3.1.3.2. Reaction Times

The reaction times were analyzed for every sentence region (Figure 10). The analyzed data only included the trials where participants had a correct response for the corresponding post-stimulus question.\textsuperscript{22} The reaction time data will be discussed in parallel with the results of the mixed effects linear regression models that examined the main effects of voice and word order, as well as the interaction effect between these two variables, for each sentence region (Table 13).

\textbf{Figure 10.} Reaction times (log-transformed) in the four conditions across the sentence regions. Error bars reflect 95\% confidence intervals.

\textsuperscript{21} Mixed effects logistic regression model used:

\texttt{glmer(Correct ~ Voice*WordOrder + (1|Participant:WordOrder) + (Voice|Item))}

\textsuperscript{22} Because the reading times for the incorrect responses tended to be aberrant and noisy, these were excluded in the data analysis. However, part of the data exploration showed that there was no major difference in the results and inferential testing when the incorrect responses were included.
Table 13. Results of the mixed effects linear regression models that tested the reaction time differences between conditions, per sentence region. The critical regions are shaded.\(^{23}\)

<table>
<thead>
<tr>
<th>Region</th>
<th>Intercept</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>Intercept</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>6.80</td>
<td>0.03</td>
<td>202.28</td>
<td>&lt; .001</td>
<td>6.98</td>
<td>0.03</td>
<td>225.72</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>-0.02</td>
<td>0.03</td>
<td>-0.89</td>
<td>0.38</td>
<td>-0.09</td>
<td>0.03</td>
<td>-2.73</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Word Order</td>
<td>0.02</td>
<td>0.03</td>
<td>0.62</td>
<td>0.54</td>
<td>0.003</td>
<td>0.03</td>
<td>0.13</td>
<td>0.90</td>
</tr>
<tr>
<td>Voice:Word Order</td>
<td>-0.07</td>
<td>0.07</td>
<td>-1.13</td>
<td>0.26</td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.44</td>
<td>0.66</td>
</tr>
<tr>
<td>Region 2</td>
<td>5.78</td>
<td>0.02</td>
<td>279.46</td>
<td>&lt; .001</td>
<td>7.04</td>
<td>0.03</td>
<td>204.89</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.64</td>
<td>0.52</td>
<td>0.02</td>
<td>0.03</td>
<td>0.79</td>
<td>0.43</td>
</tr>
<tr>
<td>Word Order</td>
<td>0.03</td>
<td>0.02</td>
<td>1.64</td>
<td>0.11</td>
<td>0.09</td>
<td>0.03</td>
<td>2.99</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Voice:Word Order</td>
<td>0.07</td>
<td>0.04</td>
<td>1.75</td>
<td>0.08</td>
<td>-0.06</td>
<td>0.06</td>
<td>-0.95</td>
<td>0.34</td>
</tr>
<tr>
<td>Region 3</td>
<td>6.49</td>
<td>0.02</td>
<td>318.83</td>
<td>&lt; .001</td>
<td>6.94</td>
<td>0.03</td>
<td>227.57</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.44</td>
<td>0.67</td>
<td>0.004</td>
<td>0.04</td>
<td>0.11</td>
<td>0.91</td>
</tr>
<tr>
<td>Word Order</td>
<td>0.02</td>
<td>0.02</td>
<td>1.23</td>
<td>0.22</td>
<td>-0.04</td>
<td>0.03</td>
<td>-1.42</td>
<td>0.16</td>
</tr>
<tr>
<td>Voice:Word Order</td>
<td>0.09</td>
<td>0.04</td>
<td>2.34</td>
<td>&lt; 0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.97</td>
<td>0.34</td>
</tr>
</tbody>
</table>

For Region 1 (first adjunct), the mean reaction times across conditions were comparable, with the exception of a slightly longer mean reaction time in the agent voice condition of the Verb-Patient-Agent word order. However, the mixed effects model for this region showed non-significant effects of Voice (\(p = 0.38\)), Word Order (\(p = 0.54\)), and their interaction (\(p = 0.26\)). Subsequently, in Region 2 (verb + adverb), there were observably shorter mean reaction times for

\(^{23}\)Mixed effects linear regression models used

\texttt{lmer(logR1 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))}
\texttt{lmer(logR2 ~ Voice*WordOrder + (Voice+WordOrder|Participant) + (Voice*WordOrder|Item))}
\texttt{lmer(logR3 ~ Voice*WordOrder + (Voice+WordOrder|Participant) + (Voice*WordOrder|Item))}
\texttt{lmer(logR4 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))}
\texttt{lmer(logR5 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))}
\texttt{lmer(logR6 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice+WordOrder|Item))}
\texttt{lmer(logR7 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))}
the two patterns in the patient voice condition compared to the agent voice condition. There was a main effect of Voice \((t = -0.89, p < 0.01)\), but no significant effect of Word Order \((t = 0.62, p = 0.90)\). No interaction effect between Voice and Word Order for this region \((t = -1.13, p = 0.66)\) was observed.

A gradience in the expected order of the mean reaction times was observed in Region 3 (nominal marker + numeral of NP1), in the order that was predicted. The mean reaction times were shortest in the patient voice condition of the Verb-Agent-Patient pattern; this was followed by comparably similar mean reaction times in the two agent condition patterns. The longest reaction times were observed in the patient voice condition of the Verb-Patient-Agent word order. Despite the observed gradience in mean reaction times across conditions, the mixed effects model showed a non-significant main effect of Voice \((t = -0.64, p = 0.52)\) and of Word Order \((t = 1.64, p = 0.11)\). The interaction effect between Voice and Word Order, which tests the predicted gradience of patterns, was non-significant, failing to support the hypothesized pattern \((t = 1.75, p = 0.08)\).

In Region 4 (adjective + nominal of NP1), there were shorter reaction times for the Verb-Agent-Patient word order patterns compared to the opposite order. The difference between these two word order patterns was significant \((t = 2.99, p < 0.01)\). There was a non-significant main effect of Voice \((t = 0.79, p = 0.43)\), and there was no interaction effect between Voice and Word Order in this region \((t = -0.95, p = 0.34)\).

A similar gradience of sentence patterns in the order predicted for Region 3 was observed in Region 5 (nominal marker + numeral of NP2). PV VAP had the shortest reaction times, while PV VPA had the longest reaction times, and the reaction times of the two agent voice patterns fell between these two patient voice patterns. In Region 5, there was a non-significant main effect of Voice \((t = -0.44, p = 0.67)\) and Word Order \((t = 1.23, p = 0.22)\). The significant interaction effect between Voice and Word Order which captures the hypothesized gradience predicted to be in Region 3, was attested in Region 5 \((t = 2.34, p < 0.05)\).

In Region 6 (adjective + nominal of NP2), there appeared to be shorter mean reaction times between the Verb-Patient-Agent conditions compared to the conditions with the opposite word order. However, this difference was non-significant and negligible. No significant effect of Word Order \((t = -1.42, p = 0.16)\) or Voice \((t = 0.11, p = 0.91)\) was observed, nor was there a significant interaction effect between Voice and Word Order \((t = 0.97, p = 0.34)\).
In Region 7 (second adjunct), no differences in mean reaction time values across conditions were observed. As shown in the mixed effects model, there was no significant effect of Voice ($t = 0.54, p = 0.59$) nor of Word Order in this region ($t = 0.09, p = 0.93$). Additionally, there was also no significant interaction effect between these two variables ($t = 0.21, p = 0.84$).

3.1.4. Discussion

The goal of Experiment 3 was to determine whether Tagalog comprehenders use voice morphology on the verb to generate expectations of the particular sentential word order, measured in terms of surprisal. Experiment 3 tested two hypotheses that examined how the expectations are shaped in real-time throughout the sentence. The first hypothesis tested whether comprehenders would generate surprisal values following the gradience of sentential patterns upon encountering the nominal marking of the first NP at Region 3, while the second hypothesis tested whether comprehenders would update their beliefs at the nominal marking of the second NP after accruing information from the verb and the first NP at Region 5.

Before turning to the experimental predictions for critical Regions 3 and 5, I first begin with interpreting the main Voice effect observed as soon as the comprehenders encounter the verb containing the voice morphology in Region 2. This significant Voice effect verified that native Tagalog speakers have a strong preference for patient voice, congruent with the results reported in previous corpus work (Cooreman et al., 1984) and in offline work involving verbs that select an animate and specific patient (N. Tanaka, 2015). This result also contributes evidence to a patient voice preference in the real-time processing of declarative sentences in Tagalog (e.g., Garcia et al., 2020).

Region 3 was hypothesized to show the effects of predictions made at the verb, where the prediction of gradient linear word order patterns would be observed as an interaction effect of Voice and Word Order. Despite the numerical differences attested in this region, the results failed to demonstrate significant prediction effects in Region 3. Additionally, no prediction effects (in the form of gradient surprisal values, or an interaction effect) were observed at the subsequent spillover region (Region 4). While the predicted gradience on surprisal did not appear at the hypothesized region (Region 3), the gradience appeared at a later sentence region (Region 5). This was demonstrated as an interaction effect between Voice and Word Order, which aligned with the
expected gradience of sentence patterns. This interaction effect at Region 5 reflects an integration effect as information from the previous regions was accrued. The fact that no significant prediction effect was observed in Region 3 suggests the lack of robust, immediate generation of expectations for sentential argument order.

Two possible interpretations could be deduced from the results at Regions 3 and 5. First, the non-significant interaction effect at Region 3, which became significant at Region 5, could suggest that predictive processes are not immediate and fast-acting, which is contrary to the claims in the literature regarding the immediacy of prediction or the generation of expectations (DeLong et al., 2014; Kuperberg & Jaeger, 2016; Kutas et al., 2011; Yan et al., 2017). There could be a tendency for prediction to be “sluggish,” such that successful predictions could be attained after a slow accumulation of activation for the generation of syntactic expectations. This line of reasoning could also apply for integration effects. The fact that differences were in fact observed in Region 5 (contra the hypothesis) implies that integration could be “sluggish” as well, wherein a gradual buildup of activation from the former regions is necessary before integrative effects could be observed.

The alternative view would be to remain faithful to the strong prediction-as-core claim, which presupposes the immediacy of predictions. This interpretation entails that the non-significant interaction effect at Region 3 suggests no effects of successful prediction were observed, and that the significant interaction effect at Region 5 was merely an integration effect. We return to the issue of the lack of robust, immediate generation of expectations in the General Discussion chapter (Chapter 4).

Region 4, the nominal region for the first noun phrase, was established as a spillover region for the previous region in the experiment. However, a different pattern of results emerged at Region 4 than in the previous region. A Word Order effect was observed in this region, with decreased reaction times for patterns with the agent-first word order, regardless of voice. This word order effect may initially be interpreted as a low-level, lexical processing advantage for agents. Because the thematic role plausibility or reversibility of nominals was not controlled in this experiment, participants may have had shorter reaction times for nominals that were more plausibly recognized as the agent. The effect observed may have been due to low-level lexical effects of certain nouns chosen as agents than as patients (Forster & Davis, 1984; McRae et al.,
1997; Mitchell et al., 1995; Trueswell et al., 1994). However, it seems more probable that this effect be interpreted as an agent-first processing advantage in comprehension, as discussed in previous research on Tagalog (Bondoc et al., 2018) and other languages (Bever, 1970; Townsend & Bever, 2001). From a processing perspective, this agent-first advantage may also support comprehenders in recognizing agents as the salient and highly accessible thematic role in event structures, or in locating the agent in the processing of the event representation (Bock & Warren, 1985; Cohn & Paczynski, 2013; Goldin-Meadow et al., 2008; Goldin-Meadow & Feldman, 1977; Kemmerer, 2012; Mauner et al., 1995; Mauner & Koenig, 2000). It may also be the case that some good-enough processing is implemented despite the noisy signal. With a good-enough processing approach adopted, comprehenders may consider applying heuristics in parsing, such as locating an agent, even without deeply processing the syntactic structure of the sentence (Ferreira, 2003; Ferreira et al., 2002, 2009). Future research could be pursued to clarify these issues.

The belief-update, or the integration of the verb and first NP information, was expected to occur prior to Region 5, the nominal marking of the second noun phrase region. This integration was expected to be demonstrated as the absence of differences between conditions. In the last two sentence regions (Regions 6 and 7), there were no main effects of Voice and Word Order, nor an interaction effect between the two.

Also noteworthy is that different patterns of results were observed from different regions in the sentence in the experiment. From the reaction time data, three different effects were observed, namely, a significant main effect of Voice in Region 2, a non-significant interaction effect between Voice and Word Order effect in Region 3, a significant main effect of Word Order effect in Region 4, and a significant interaction effect between Voice and Word Order effect in Region 5. These different patterns highlight how the processing of voice unfolds throughout the sentence. The reaction time measures show how the expectations are shaped and modified, from the time comprehenders receive the verb and voice information, until they accrue more information about the argument structure.

Additionally, these reaction time patterns can be compared with the findings from the post-stimulus task, which showed a main effect of Word Order. The findings between the reaction time and comprehension accuracy illustrate a mismatch between the offline comprehension of the sentence and the real-time processing of the unfolding sentence (in particular, in Region 5). Similar
findings of a mismatch have been also reported in previous research on syntactic ambiguities, in which comprehenders were observed to retain interpretations that reflected incorrect syntactic parses, even though reaction time data reflected syntactic reanalyses (Christianson et al., 2001, 2006; Patson et al., 2009). I further discuss the implications of these findings relating to the nature of comprehension processes in the General Discussion chapter.

Some scholars suggest that the outcomes obtained in a self-paced reading experiment come from rough comprehension measures. But regardless of the limitations of using self-paced reading as a method, the results obtained are informative of the dynamic processing of voice throughout the sentence.

To summarize, Experiment 3 tested whether Tagalog comprehenders generated expectations for detailed sentential argument order using voice information using self-paced reading, measured in terms of surprisal. The results showed that comprehenders demonstrated the effects of hypothesized gradience of sentential word order patterns at a later sentence region, suggesting the lack of robust, immediate effects of prediction. The next section explores whether the generation of expectations is limited under cognitively constrained contexts, which would further illuminate the nature of predictive processes.

3.2. Experiment 4: Rapid Serial Visual Presentation

3.2.1. Introduction and Hypotheses
The pervasiveness of predictive processing in the cognitive system is a controversial topic in psycholinguistics. Some research claims that prediction is the core and unifying principle of the human language processing system, with some authors describing human brains as being essentially “prediction machines” (Bar, 2009; Clark, 2013; Friston, 2010). The view that prediction exists in comprehension process has been demonstrated by extensive work that utilized a variety of behavioral and neuropsychological methods (e.g., Altmann & Kamide, 1999; Chambers, Tanenhaus, & Magnuson, 2004; DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999; Tanenhaus, Eberhard, & Sedivy, 2002).

However, some scholars have also expressed reluctance to accept prediction as a pervasive process in information processing, claiming that most research that offers evidence on prediction
had experimental set-ups and stimuli items that were “prediction-encouraging” (Huettig, 2015; Huettig & Guerra, 2019). Other scholars also question the N400 component as being purely attributable to prediction effects or to integration effects (Huettig & Guerra, 2019; Nieuwland, Barr, et al., 2018; Van Petten & Luka, 2012).

The controversies on prediction has brought to bear the question on the prevalence of prediction, in other words, prediction as an active, routine operation that is constantly being generated or performed by the system (Kuperberg & Jaeger, 2016; Yan et al., 2017). If predictions or syntactic expectations are routinely being generated by the system, it would be logical to posit the prevalence of these operations across various situations, even in cognitively constrained contexts. An example of a cognitively constrained context would be conditions in which we comprehend language under time pressure. Such conditions often reflect language processes that can occur in everyday situations, given that people often engage in language activities while also engaging some other activity or task.

These cognitively constrained conditions can establish the ground from which we can test the hypothesis that predictive processes routinely operate in the system. While this provides us with an opportunity to test the hypothesis, the additional constraint on the system also impacts comprehension outcomes. An external factor, such as time pressure, may impede the accurate parsing of syntactic structure (Bradley & Shapiro, 2004; Menzel, 1994). A trade-off between speed and accuracy is also reported in the literature on speed reading (Rayner et al., 2016). But regardless of this trade-off, time-constrained conditions would be worthwhile to determine whether the same gradience of syntactic expectations prevail in cognitively constrained conditions.

Experiment 4 was a parallel study done in conjunction with the self-paced reading study (Experiment 3), where the same stimuli items were presented at two presentation rates. One version of the experiment (Experiment 4a) presented each sentence region for 500ms, while another version (Experiment 4b) presented each sentence region for 750ms. I hypothesize that the processing system still generates syntactic expectations, following the predicted gradience of linear word order, even in conditions with time pressure. The following specific predictions were tested in this experiment.
**Hypothesis 1:** There will be a speed-accuracy trade-off given the rapid presentation of the linguistic stimuli. Overall, there would be lower comprehension accuracy values in the task with greater time pressure compared to a task with less time pressure. More specifically, accuracy performance in Experiment 4a (500ms) would be lower compared to Experiment 4b (750ms). The RSVP sub-experiments would also be worse in accuracy performance compared to the self-paced reading task (Experiment 3). Hence, I predict a main effect of Task.

**Hypothesis 2:** Despite the time pressure, the system will still generate syntactic expectations as comprehenders encounter the verb and voice information, with the working assumption that prediction is a routine, immediate process that the system performs.

These probabilistic expectations will be congruent to the cline of linear word order patterns, and the result of these expectations will be reflected in the comprehension accuracy rates. Hence, comprehenders will commit the fewest errors in the strongly expected pattern (i.e., patient voice in the Agent-Patient order), and commit the most errors in the least expected pattern (i.e., patient voice in the Patient-Agent order). Meanwhile, the error rates in the two agent patterns will not differ between each other, and the agent voice rates will fall between the error rates in the patient voice.

In addition, the accuracy rates will be modulated in the conditions with time pressure, compared to the accuracy rates in the self-paced reading experiment. I predict an overall disordinal interaction effect between Voice and Word Order and a superadditive interaction effect among Voice, Word Order, and Task.

The self-paced reading study (Experiment 3) was executed in parallel with Experiment 4, therefore the predictions on the generation of graded syntactic expectations for both experiments were similar.

### 3.2.2. Method

This study was conducted using rapid serial visual presentation (RSVP), non-cumulative moving window, using the program E-Prime (Schneider et al., 2002) on a laptop.
3.2.2.1. Participants
Participants in Experiment 4 included 114 Tagalog speakers from the Metropolitan Manila region in the Philippines (age range 18-60, mean 25.85; 75 females, 39 males). In a language background questionnaire, all participants identified themselves as Tagalog-dominant. All participants also spoke English as a second language, in varying degrees of proficiency. Some participants also identified themselves as speaking a third or fourth language, but of lower proficiency compared to Tagalog and English. Regarding the Tagalog dialects spoken by participants, 93 identified as speakers of the Metro Manila dialect only, four as speakers of a northern dialect, nine as speakers of a southern dialect, and eight as speakers of the Metro Manila dialect and another northern or southern dialect. Fifty-eight participants completed Experiment 4a, and 56 participants completed Experiment 4b. None of these participants were linguistics majors.

3.2.2.2. Materials
To allow for a direct comparison with the self-paced reading task, the stimuli items and experiment lists were similar to Experiment 3, except that the stimulus onset asynchronies (SOAs) of the sentence regions were manipulated. Experiment 4a presented each sentence region for 500ms SOA, while Experiment 4b presented the sentence regions at 750ms SOA. A similar post-stimulus task (yes/no questions) to that of Experiment 3 appeared after the sentence presentation.

3.2.2.3. Procedure
Participants were assigned in a rotating fashion to one of the lists. They were asked to press a button to begin the trial. Afterward, participants responded to the post-stimulus task (yes/no comprehension question) using keyboard buttons.

3.2.2.4. Analyses
A comparison between the accuracy of the responses between the two RSVP tasks and the post-stimulus task in self-paced reading was conducted using a maximal mixed effects logistic regression model (Barr et al., 2013). Voice, Word Order, and Task were treated as deviation-coded fixed effects, while participants and items were treated as random effects. The intercepts and slopes for the fixed effects were initially included in the random effects structure. Because of non-
convergence, the model was simplified. The final model had Voice and Task nested for Participant, and the intercept and slope for Voice was included for Item.

3.2.3. Results
A decline was observed in offline comprehension accuracy rates in the two RSVP tasks (the 500ms and 750ms panels of Figure 11) compared to the self-paced reading task. Performance between the two RSVP tasks was comparable. In terms of the sentential patterns, there were higher accuracy rates in the agent-patient word order patterns in the self-paced reading task compared to the patient-agent patterns; this word order difference diminished in the two RSVP sub-experiments. Furthermore, in both of the RSVP tasks, accuracy was the highest for the highly preferred agent-patient word order in the patient voice, while there were no differences in accuracy between the other patterns.

![Figure 11](image-url)

**Figure 11.** Accuracy rates in the two voice and word order conditions in Experiments 4a (500ms), 4b (750ms) and the self-paced reading task. Error bars reflect 95% confidence intervals.

Because of the similarity of the results between Experiments 4a and 4b, these conditions were recategorized under a single factor (“RSVP condition”) in the statistical analyses. The mixed
effects logistic regression model demonstrated a significant main effect of Task ($z = 8.79$, $p < 0.001$), suggesting diminishing accuracy from self-paced reading to the RSVP task. A significant main effect of Word Order ($z = -7.12$, $p < 0.001$) was observed, pointing toward an overall advantage to sentences with the Agent-Patient word order. There was also a significant interaction between Word Order and Task ($z = -5.31$, $p < 0.001$), which showed that the Word Order effect weakened in the RSVP task. A significant interaction between Voice and Word Order ($z = -2.54$, $p < 0.01$) was observed, which highlighted the highest accuracy rate for the most frequent pattern (PV VAP). No other main and interaction effects were significant.

Table 14. Results of the mixed effects logistic regression models that examined comprehension accuracy between the two RSVP conditions and the post-stimulus task in self-paced reading.\footnote{Mixed effects logistic regression models used \texttt{glmer(Correct ~ Voice*WordOrder*Task + (1|Participant:Voice) + (Voice|Item) + (1|Participant:Task))}}

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>SE</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.17</td>
<td>0.10</td>
<td>11.51</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>0.32</td>
<td>0.17</td>
<td>1.83</td>
<td>0.07</td>
</tr>
<tr>
<td>Word Order</td>
<td>-0.76</td>
<td>0.11</td>
<td>-7.12</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Task</td>
<td>0.98</td>
<td>0.11</td>
<td>8.79</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice:Word Order</td>
<td>-0.54</td>
<td>0.21</td>
<td>-2.54</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Voice:Task</td>
<td>-0.17</td>
<td>0.13</td>
<td>-1.28</td>
<td>0.20</td>
</tr>
<tr>
<td>Word Order:Task</td>
<td>-0.71</td>
<td>0.13</td>
<td>-5.31</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice:Word Order:Task</td>
<td>0.07</td>
<td>0.27</td>
<td>0.28</td>
<td>0.78</td>
</tr>
</tbody>
</table>

3.2.4. Discussion

The aim of Experiment 4 was to examine whether predictive processing or the generation of expectations, often construed as an immediate and fast-acting process, would regularly manifest, even in contexts with time pressure. This question was investigated using RSVP, where sentential regions were presented with either 500ms or 750ms SOA. I hypothesized (i) a speed-accuracy trade-off between the RSVP tasks and the self-paced reading task, and (ii) the generation of expectations, following the predicted gradience of sentential word order patterns, in each task.

The experimental data showed that the first hypothesis was met, confirmed by a main effect of Task. This is in line with research that shows comprehension processes in reading are affected as
a function of speed, where comprehension declines given reduced time for processing (Juola et al., 1982; Potter et al., 1980; Rayner et al., 2016; Schotter et al., 2014).

The results show that the second hypothesis was not supported, as the predicted gradience of word order patterns was not observed in either task. However, a significant overall main effect of Word Order was observed, pointing toward an agent-first advantage. This Word Order effect, which was clearly demonstrated in the self-paced reading condition, diminished in the RSVP condition. Such weakening of the word order advantage in the RSVP condition was captured as a significant interaction between Word Order and Task. This agent-patient word order advantage can be explained by the good-enough processing that comprehenders do when understanding a sentence, whereby comprehenders tend to rely on heuristics such as identifying the agent without deeply processing the syntactic structure (as discussed in Section 3.1.4, Ferreira, 2003; Ferreira et al., 2002). However, the diminishing Word Order effect in the RSVP tasks is suggestive that it is difficult for comprehenders to even compute good-enough processing across the different sentential patterns under cognitively constrained contexts. It is important to note that the comprehension questions in Experiment 3 tested the thematic roles of the nominals in the event (i.e., the who-did-what of the event; see Methods Section 3.1.2.3 regarding the construction of the comprehension questions). With the restricted time to process the unfolding material, comprehenders found it difficult to make successful predictions and integrate or update the thematic roles of the event participants in the developing event representation.

These findings can also be explained with recent models that incorporate the notion of noise in sentence processing. In noisy-context models, the comprehender is faced with the challenge of accurately extracting the information despite the noise in the signal (Bergen et al., 2012; Futrell & Levy, 2017; Gibson et al., 2013). Time pressure may be described as inducing noise in the signal, thereby affecting the accuracy of the comprehenders’ parse of the sentence. In the present experiment, time pressure may be likened to noise. From a noisy-context perspective, comprehenders may interpret the signal as problematic and may develop a syntactic “correction” of what they assumed to be the likely structure of the signal, i.e., the highly expected patient voice in the Agent-Patient word order. In lossy-context models, the fast presentation of the sentence creates a lossy memory representation of the event (Futrell et al., 2020). In either of these scenarios, the noise impact leads the comprehender to apply heuristics in decoding the signal or in generating.
the lossy event representation. Applying the concept of lossy-context processing in the present experiment, the comprehenders may produce an incorrect parse of the signal and generate a corrupted memory representation of the event, resulting to an error, such as a thematic role switching of the nominal entities in the event. These points are discussed further in Chapter 4.

Across all tasks in this experiment, participants had the highest accuracy scores in the strongly expected patient voice Agent-Patient word order, while the expected gradience in the other patterns was not observed for the other three patterns. This overall high accuracy in the strongly expected pattern was shown as a significant interaction between Voice and Word Order. This result is congruent with previous findings that described the impact of stimulus presentation timing on predictability effects in comprehension (Masson, 1986; Stanovich & West, 1979; Traxler & Foss, 2000). In particular, comprehenders tend to make successful predictions only in highly predictable (high-cloze) contexts within the context of speeded presentation rates (Chow et al., 2018; Dambacher et al., 2012; Ito et al., 2016; Luka & Van Petten, 2014; Wlotko & Federmeier, 2015), which aligns with Experiment 4 findings. This suggests that gradience in predictive processing is highly reduced (e.g., medium-cloze vs. low-cloze) when processing in cognitively constrained contexts. In order for comprehenders to make a successful prediction, the comprehender would be expected to simultaneously perform multiple operations, such as the uptake of information, generation of expectations based on accrued input, integration of newly accrued input, and belief-updating. This comprehension failure may be related to the system’s limited resources to simultaneously perform all functions, with the accrued information “buffering” in working memory while there is continued uptake of input (Poldrack et al., 2001; Vagharchakian et al., 2012). The relationship between prediction, other comprehension processes, and memory constraints warrants future investigation.

The experimental findings also did not show a difference between 500ms and 750ms. It is common in sentence processing research to employ a word-per-word presentation rate of 700ms to measure effects of prediction (e.g., Dambacher et al., 2012; Hill et al., 2005; Luka & Van Petten, 2014); however, it is critical to point out two major differences in this study from other experiments. In contrast to other research which implements sentence presentation on a word-by-word basis, this study presented stimuli by regions which contained two to three words, thereby necessitating more time to process each region. The participants in the self-paced reading study
had reaction time values of at least 600ms for the regions with function words, such as nominal marking (mean RT of R3 = 767ms; mean RT of R5 = 697ms) and had at least 1000ms for those that contained major content words, such as the verb and the nominals (mean RT of R2 = 1172ms; mean RT of R4 = 1248ms; mean RT of R6 = 1137ms). Given the amount of time that comprehenders actually tended to take to process the stimuli, it is reasonable to find a lack of difference between the 500ms and 750ms RSVP sub-experiments.\(^{25}\) Even the sub-experiment with the longer time (750ms) resulted in substantial time pressure; therefore, both sub-experiments can be considered as equally difficult settings for sentence parsing. In both sub-experiments, the information was presented too rapidly for comprehenders to simultaneously perform multiple operations that would result in successful comprehension. Future research could explore whether RSVP manipulations which would be similar to the self-paced reading rates would show similar results.

The overall results of Experiment 4 suggest the limited effects of prediction from voice information in sentence processing. In particular, the word order advantage observed in Experiment 3 diminished in Experiment 4, and only an effect of the highly preferred form (the patient voice of the Verb-Agent-Patient pattern) was observed. The results of the experiment also reflect the limited effects of integration from voice and argument information in cognitively constrained contexts and reveal two possible insights on predictive processing: either predictions can be a slowly accumulating process and thus would not ubiquitously result to successful outcomes, or predictions could be immediate and fast-acting, but successful predictions are only obtained from those patterns with higher probability. I discuss this point further and how this relates to the nature of predictive processes in the General Discussion chapter (Chapter 4).

\(^{25}\)Note that the RSVP experiment (Experiment 4) was implemented in parallel with the self-paced reading experiment (Experiment 3); hence, Experiment 3 findings were not factored into the RSVP manipulations.
3.3. **Experiment 5: Self-Paced Reading with Memory Load**

3.3.1. Introduction and Hypotheses

The findings from Experiment 4 showed that participants did not have comprehension accuracy rates that patterned with the predicted cline of linear word order patterns; however, the experiment did not investigate the process by which expectations are generated in real-time in cognitively constrained contexts. Experiment 5 investigates whether prediction is a fundamental process that comprehenders routinely engage in real-time to anticipate the unfolding linguistic signal in cognitively constrained contexts. In Experiment 5, I attempt to address this question directly by implementing the same self-paced reading experiment in Experiment 3 with increased memory load.

The imposition of a cognitive constraint, such as memory load, allows us to test the hypothesis that prediction is a core principle of the processing system. The Prediction-as-Core Hypothesis presumes that prediction, or the active generation of expectations, is a fundamental, routine process of the processing system. This hypothesis speculates that syntactic expectations would be generated and used in real-time despite the presence of an additional constraint. Many recent accounts of prediction support this hypothesis (Altmann & Mirković, 2009; Dell & Chang, 2013; DeLong et al., 2014; Kuperberg & Jaeger, 2016; Pickering & Garrod, 2007, 2013).

One of the ways an additional constraint like memory load can be imposed on the information processing system in an experiment would be to require the comprehender to actively retain information in working memory, such as a series of letters while another cognitive operation (i.e., parsing of the linguistic signal) is ongoing. With the assumption of a single resource pool for information processing (Baddeley et al., 1987; Just & Carpenter, 1992), the simultaneous performance on tasks could reduce the processing resources that could efficiently be used in interpreting the linguistic signal.

Using the same stimuli in the self-paced reading study, this experiment tests for the routine implementation of prediction in language comprehension. The following hypotheses on predictive processing were tested by comparing the outcomes in this experiment against the results in Experiment 3.
**Hypothesis 1:** Comprehension accuracy will diminish in the self-paced reading task that includes the memory load, in comparison to the task that excludes it, given the tradeoff between accuracy and memory load. I predict a main effect of Experiment on comprehension accuracy.

**Hypothesis 2:** Across the board, reaction times will be elevated in the self-paced reading task with the memory load, compared to the task that excludes it. I predict a main effect of Experiment on reaction times.

**Hypothesis 3:** At the nominal marking of the first nominal argument (i.e., Region 3), a gradience in reaction times will be observed, following the proposed hierarchy of syntactic expectations in Experiment 3.

The Prediction-as-Core Hypothesis theorizes the active and immediate generation of expectations is a crucial part of the information processing system; hence, predictions shall be routinely generated even in an experiment with an induced cognitive load. If the Prediction-as-Core Hypothesis holds, regardless of the imposition of memory load, the probabilistic syntactic expectations for the different sentence patterns would still be observed. Within this experiment, patient voice in the Agent-Patient order would obtain the lowest RT (the least surprisal), while patient voice in the Patient-Agent order would have the highest RT (the most surprisal). Both patterns in the agent voice would have RTs or surprisal values in between the two patient voice patterns.

Furthermore, due to the imposition of the memory load, comprehenders are expected to have magnified RTs in this task compared to Experiment 3. While the mean reaction times in Experiment 5 follow the predicted cline of sentential patterns, there will be a superadditive increase in the RT of each sentential pattern in Experiment 5 compared to its counterpart in Experiment 3. The addition of memory load makes the task more difficult; this increase in difficulty would also further illuminate the differences between conditions in Region 3.

Thus, I predict a main effect of Experiment, an overall disordinal interaction effect between Voice and Word Order, and a superadditive interaction effect between Voice, Word Order, and Experiment, at Region 3.
**Hypothesis 4:** At Region 5, there will be a significant interaction effect between Voice and Word Order for both experiments, as was observed in Experiment 3. At the same time, there will also be a superadditive elevation in reaction times in Experiment 5 compared to Experiment 3. There would still be a main effect of Experiment, a disordinal interaction between Voice and Word Order, and a superadditive interaction effect between Voice, Word Order, and Experiment.

The addition of the memory load highlights the differences between conditions. The predicted cline of surprisal values is hypothesized to occur in Region 5. At this point in the sentence, the system has accrued adequate information from the verb and the first noun phrase, which leads to continued, gradient activation of the likely word order patterns. As a result, the ease of integrating the unfolding material should follow the possible predicted cline of surprisal values from Region 3. Unlike the proposed hypothesis of a belief update in Experiment 3, the cognitive system may need more time, or more information, such as the upcoming second argument, before it engages in belief updating.

### 3.3.2. Method

To test the study hypotheses, the self-paced reading (non-cumulative moving window) task from Experiment 3 was implemented using IbexFarm ver 0.3.9 (Drummond, 2013). To induce memory load while reading sentences, participants were asked to remember a series of three letters (such as Z M X) that were flashed for one second prior to sentence presentation. Participants were asked to recall these letters after responding to the comprehension question.
3.3.2.1. Participants

Ninety-seven native Tagalog speakers from the Philippines (age range 18-52, mean age 27.38; 77 females) were recruited online for Experiment 5. Out of the 97 participants, 71 identified themselves as speakers of Metro Manila Tagalog, 20 identified themselves as speakers of a Southern Tagalog dialect, and 6 of a Northern Tagalog dialect. None of the participants had any visual or hearing issues, or a history of a communication disorder. In a language background questionnaire, all participants identified themselves as Tagalog-dominant, and as second language English speakers with varying degrees of proficiency. Some participants also identified themselves as speaking a third or fourth language, but of lower proficiency compared to Tagalog and English.

3.3.2.2. Materials

To allow for direct comparison between experiments, all sentence stimuli, comprehension questions, and experimental lists were identical to Experiments 3 and 4, with the exception of three filler items that were modified for each list with sentences involving exclamatives in Region 1, as a counter-check for unusually long reading times from participants. For the memory load task, a series of three letters was pseudo-randomly generated for every trial. Each series of letters in the memory load only contained consonants. No two consecutive trials had a similar letter in the memory load.

3.3.2.3. Procedure

Participants were assigned in a rotating fashion to one of the lists. Prior to every sentence presentation, the series of letters was flashed on the screen for one second. Participants pressed a key to read each sentence fragment. Afterwards, participants responded to the post-stimulus task (yes/no question) by either clicking on the Yes/No options on the screen or by doing keyboard presses. The trial ended with the participants asked to recall and type in a textbox the series of letters flashed prior to sentence presentation.
3.3.2.4. Analyses

To determine whether there were differences in memory load performance as a function of voice and word order pattern in Experiment 5, a maximal mixed effects logistic regression model was made (Barr et al., 2013), with Voice and Word Order as fixed effects, and participants and items as random effects. Intercepts and slopes of the fixed effects were included in the random effects structure, with deviation coding applied to both fixed effects. The model was simplified by removing terms one by one if the model did not converge. Only those memory load responses with all three letters correct and in the right order were considered as accurate.

To allow for a direct comparison between the performance in comprehension accuracy and reaction times between Experiments 3 and 5, the data sets from these two experiments (post-exclusion of outlier participants) were combined for statistical analyses. A comparison between the accuracy of the poststimulus task responses between Experiments 3 and 5 was made, using a maximal mixed effects logistic regression model, with Experiment, Voice, and Word Order as deviation-coded fixed effects, and participants and items as random effects. Intercepts and slopes of the fixed effects were included in the random effects structure. A similar strategy of model simplification, described above, was applied to allow for model convergence. For both experiments, only those trials with correct comprehension accuracy responses were counted for reaction time data analyses.26

The reaction time data after the exclusion of outlier participants and trials for Experiments 3 and 5 were also combined.27 To examine whether the same gradient syntactic expectations were generated between Experiments 3 and 5, critical comparisons were made with the reaction times in all sentence regions between and across conditions and experiments. The reaction times were analyzed using a maximal mixed effects linear regression model in R (R Core Team, 2013) using the lme4 package (Bates et al., 2014), with the reaction times as the dependent variable, Voice, Word Order, and Experiment as predictors, and with items and participants as random effects. Intercepts and slopes of the fixed effects were also incorporated in the random effects structure.

26A subset analysis was implemented, which included trials with accurate responses in both comprehension accuracy and memory load. Due to significant data loss, this analysis was not chosen for statistical analyses.  
27Readers who would like to look at an isolated statistical analysis of Experiment 5 alone can refer to Appendix F.
with the same model simplification applied in cases of non-convergence. As in the other models, deviation coding was applied to obtain the desired main and interaction effects.

3.3.3. Results
Nine participants were removed from the analysis due to a low comprehension accuracy score and/or a low memory task score (< 60%), leaving data from 88 participants for further analysis. This section presents the results from the comprehension and memory tasks, followed by the region-by-region reaction time results, after the exclusion of these participants.

3.3.3.1. Memory Load Accuracy
No difference in the accuracy of participants in the memory load task was observed regardless of the condition in Experiment 5 (Figure 13). The accuracy rates ranged from 80.88 – 83.20% (sd range: 0.37 – 0.39) across the different sentence patterns. This lack of a difference was also supported by the non-significant main effect of Voice ($z = 0.04$, $p = 0.97$) and of Word Order ($z = -1.65$, $p = 0.10$), as well as a lack of interaction effect between Voice and Word Order ($z = 0.24$, $p = 0.81$).

![Figure 13. Accuracy rates in the memory load task across conditions in Experiment 5. Error bars reflect 95% CI.](image)
Table 15. Results of the mixed effects logistic regression models for memory load accuracy.\textsuperscript{28}

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>SE</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.71</td>
<td>0.10</td>
<td>16.58</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>0.003</td>
<td>0.10</td>
<td>0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>WordOrder</td>
<td>-0.19</td>
<td>0.11</td>
<td>-1.65</td>
<td>0.10</td>
</tr>
<tr>
<td>Voice:WordOrder</td>
<td>0.05</td>
<td>0.19</td>
<td>0.24</td>
<td>0.81</td>
</tr>
</tbody>
</table>

3.3.3.2. Comprehension Accuracy

As shown in Figure 14, the accuracy rates of participants between Experiments 3 and 5 were comparable. Overall, there was a slightly higher accuracy for questions involving the patient voice ($\bar{x} = 78.28\%$, $sd = 0.41$) compared to the agent voice ($\bar{x} = 76.33\%$, $sd = 0.43$); however, the difference between these two voice conditions was not significant ($z = 1.18$, $p = 0.24$). There was an observable word order effect for both experiments. Participants had better performance in the Agent-Patient patterns ($\bar{x} = 85.61\%$, $sd = 0.35$) compared to the Patient-Agent patterns ($\bar{x} = 69.01\%$, $sd = 0.46$), and the difference between these two word order patterns was significant ($z = -15.51$, $p < 0.001$). Accuracy rates in Experiment 3 appeared to be slightly better ($\bar{x} = 79.13\%$, $sd = 0.41$) than in Experiment 5 ($\bar{x} = 75.87\%$, $sd = 0.43$), but the difference between the two experiments was not significant ($z = -1.60$, $p = 0.11$).

\textbf{Figure 14.} Comprehension accuracy rates in the different voice and word order conditions between Experiments 3 and 5. Error bars reflect 95% CI.

\textsuperscript{28} Mixed effects logistic regression model used
\texttt{glmer(MemoryScore ~ Voice*WordOrder + (WordOrder|Participant) + (1|ItemNumber:Voice), data, family="binomial")}
In terms of the cross between voice and word order patterns, participants had the overall highest accuracy rate for patient voice condition of the Verb-Agent-Patient pattern ($\bar{x} = 87.01\%, sd = 0.34$), followed by the agent voice condition of the same word order ($\bar{x} = 84.20\%, sd = 0.36$). Accuracy in the other patterns involving the opposite word order, namely, Verb-Patient-Agent pattern in the patient voice ($\bar{x} = 69.55\%, sd = 0.46$) and agent voice ($\bar{x} = 68.47\%, sd = 0.46$) were low. No significant interaction effect was observed between Voice and Word Order ($z = -1.28, p = 0.20$), Voice and Experiment ($z = -0.99, p = 0.32$), and Word Order and Experiment ($z = -0.24, p = 0.81$). A three-way interaction between Voice, Word Order, and Experiment was also not observed ($z = 0.47, p = 0.64$). Table 16 summarizes the results of the mixed-effects logistic regression model.

**Table 16. Results of the mixed effects logistic regression models for comprehension accuracy.**

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>$SE$</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.69</td>
<td>0.19</td>
<td>8.82</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>0.12</td>
<td>0.10</td>
<td>1.18</td>
<td>0.24</td>
</tr>
<tr>
<td>WordOrder</td>
<td>-1.19</td>
<td>0.08</td>
<td>-15.51</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Experiment</td>
<td>-0.22</td>
<td>0.13</td>
<td>-1.60</td>
<td>0.11</td>
</tr>
<tr>
<td>Voice:WordOrder</td>
<td>-0.19</td>
<td>0.15</td>
<td>-1.28</td>
<td>0.20</td>
</tr>
<tr>
<td>Voice:Experiment</td>
<td>-0.15</td>
<td>0.15</td>
<td>-0.99</td>
<td>0.32</td>
</tr>
<tr>
<td>WordOrder:Experiment</td>
<td>-0.04</td>
<td>0.15</td>
<td>-0.24</td>
<td>0.81</td>
</tr>
<tr>
<td>Voice:WordOrder:Experiment</td>
<td>0.14</td>
<td>0.30</td>
<td>0.47</td>
<td>0.64</td>
</tr>
</tbody>
</table>

### 3.3.3.3. Reaction Times

Similar to Experiment 3, the reaction times were analyzed for every sentence region (Figure 15). The data from Experiment 3 and 5 only included the trials where participants had a correct response for the corresponding comprehension question. An additional analysis involving the exclusion of data with incorrect responses to the memory load task in Experiment 5 was also explored. The results were similar to the data set that included it. However, to avoid significant data loss, I decided to include trials with incorrect responses to the memory load task.

---

29 Mixed effects logistic regression model used

```r
glmer(Score ~ Voice*WordOrder*Experiment + (1|Participant) + (Voice|ItemNumber), data, family="binomial")
```

30 An additional analysis involving the exclusion of data with incorrect responses to the memory load task in Experiment 5 was also explored. The results were similar to the data set that included it. However, to avoid significant data loss, I decided to include trials with incorrect responses to the memory load task.
In Region 1 (first adjunct), the mean reaction times across the different sentence patterns were comparable, with a slight elevation of reaction times in Experiment 5 versus Experiment 3. The mixed effects model showed a significant main effect of Experiment ($t = 4.01, p < .001$), as well as a significant interaction effect between Voice and Experiment ($t = 2.03, p < .05$). There was no main effect of Voice ($t = 0.48, p = 0.63$) or Word Order ($t = 1.67, p = 0.10$) observed in this region, nor was there an interaction between Voice and Word Order ($t = -1.54, p = 0.12$), Word Order and Experiment ($t = 0.60, p = 0.55$), or Voice, Word Order, and Experiment ($t = 1.39, p = 0.66$).

An asymmetry between the two voice conditions was observed in Region 2 (verb + adverb). For both experiments, the reaction times for patient voice was shorter than agent voice. The mixed effects model for this region only showed a significant main effect of Voice ($t = -3.75, p < .001$). There was no significant main effect of Word Order ($t = -0.77, p = 0.45$) or Experiment ($t = 1.09, p = 0.27$) in this region, nor was there a significant interaction effect between Voice and Word Order ($t = 1.70, p = 0.09$), Voice and Experiment ($t = 0.98, p = 0.33$), Word Order and Experiment ($t = 1.12, p = 0.26$), or Voice, Word Order, and Experiment ($t = 0.38, p = 0.70$).

The gradience in the expected order of the mean reaction times observed in Region 3 (nominal marker + numeral of NP1) in Experiment 3 diminished in Experiment 5. In Experiment 3, the mean reaction time was the shortest for patient voice in the Verb-Agent-Patient order, the longest in the patient voice pattern in the Verb-Patient-Agent order. The mean reaction times of the two agent
voice patterns was in between the mean reaction time of the two patient voice patterns. The predicted gradient pattern was not observed, and the reaction times were similar across conditions in Experiment 5. Interestingly, unlike in Experiment 3, the mixed effects model involving the data from both experiments showed a significant interaction effect of Voice and Word Order ($t = 2.51, p = 0.01$) in this region, supporting the hypothesized gradience of sentence patterns. No significant main effects were observed for Voice ($t = 0.64, p = 0.53$), Word Order ($t = 0.77, p = 0.45$) or Experiment ($t = 0.41, p = 0.69$) in Region 3, nor was there a significant interaction effect between Voice and Experiment ($t = 1.53, p = 0.13$), Word Order and Experiment ($t = -1.65, p = 0.10$), and Voice, Word Order, and Experiment ($t = -0.22, p = 0.82$).

Experiment 3 demonstrated a sharper distinction between the two word order patterns in Region 4 (adjective + nominal of NP1) than in Experiment 5; in other words, in Experiment 3, comprehenders exhibited shorter reaction times for the Verb-Agent-Patient word order compared to the opposite order. Although the mean reaction times between the two word orders was maintained in Experiment 5, the difference was less prominent. This word order difference was not supported by the mixed effects model, which did not show a significant main effect of Word Order ($t = 1.62, p = 0.12$). Furthermore, the model also did not report a significant main effect of Voice ($t = -0.77, p = 0.45$) or Experiment ($t = -1.06, p = 0.29$) in this region. There were also no significant interaction effects between Voice and Word Order ($t = -0.81, p = 0.42$), Voice and Experiment ($t = -0.31, p = 0.76$), Word Order and Experiment ($t = -1.86, p = 0.06$), or Voice, Word Order, and Experiment ($t = -0.35, p = 0.72$).

In Experiment 3, the expected gradience of sentence patterns in the predicted order for Region 3 was observed for Region 5 (nominal marker + numeral of NP2). This gradient pattern was less distinctive in Experiment 5. Despite the appearance of reduction in the gradience, the mixed effects model supports the hypothesized gradience with a robust interaction effect between Voice and Word Order ($t = 4.15, p < .001$). In Region 5, there was no significant main effect of Voice ($t = 0.61, p = 0.54$), Word Order ($t = 0.93, p = 0.36$) or of Experiment ($t = 1.91, p = 0.06$), nor was there a significant interaction effect between Voice and Experiment ($t = 0.20, p = 0.84$), Word Order and Experiment ($t = -0.62, p = 0.53$), or Voice, Word Order, and Experiment ($t = -0.45, p = 0.65$).
In Region 6 (adjective + nominal of NP2), the mean reaction time for the Verb-Patient-Agent conditions compared to the Verb-Agent-Patient conditions was shorter for both experiments. However, this difference was not significant, as shown by a non-significant main effect of Word Order in this region \( (t = -1.78, p = 0.08) \). There was also no significant main effect of Voice \( (t = -0.08, p = 0.94) \) or Experiment \( (t = -0.73, p = 0.47) \) in this region, nor was there a significant interaction effect between Voice and Word Order \( (t = 0.74, p = 0.46) \), Voice and Experiment \( (t = -0.13, p = 0.90) \), Word Order and Experiment \( (t = 0.03, p = 0.97) \), or Voice, Word Order, and Experiment \( (t = -0.88, p = 0.38) \).

In the final region, Region 7 (second adjunct), the mean reaction times for the different sentence patterns appeared to not differ between Experiment 3 and Experiment 5. The apparent lack of a difference between conditions is demonstrated by the results of the mixed effects model, which showed a non-significant main effect of Voice \( (t = 0.67, p = 0.51) \), Word Order \( (t = -0.44, p = 0.66) \), and Experiment \( (t = 0.37, p = 0.71) \), as well as a non-significant interaction effect between Voice and Word Order \( (t = 1.33, p = 0.18) \), Voice and Experiment \( (t = -0.46, p = 0.65) \), and Word Order and Experiment \( (t = 0.34, p = 0.74) \), and Voice, Word Order, and Experiment \( (t = 1.29, p = 0.20) \).
Table 17. Results of the mixed effects linear regression models that tested the reaction time differences between conditions, per sentence region. The critical regions are shaded.\[^{31}\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.88</td>
<td>0.03</td>
<td>0.03</td>
<td>0.18</td>
<td>0.01</td>
<td>-0.04</td>
<td>0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>220.27</td>
<td>0.48</td>
<td>1.67</td>
<td>4.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.77</td>
</tr>
<tr>
<td></td>
<td>&lt; .001</td>
<td>0.63</td>
<td>0.10</td>
<td>&lt; .001</td>
<td>0.63</td>
<td>0.12</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>2</td>
<td>6.99</td>
<td>0.03</td>
<td>0.03</td>
<td>0.05</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>241.61</td>
<td>0.08</td>
<td>0.77</td>
<td>2.03</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.38</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>&lt; .001</td>
<td>0.45</td>
<td>1.70</td>
<td>&lt; .05</td>
<td>-0.01</td>
<td>-0.31</td>
<td>0.70</td>
<td>0.26</td>
</tr>
<tr>
<td>3</td>
<td>6.57</td>
<td>0.02</td>
<td>0.01</td>
<td>0.06</td>
<td>0.03</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.35</td>
</tr>
<tr>
<td></td>
<td>372.27</td>
<td>0.53</td>
<td>0.45</td>
<td>2.51</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.05</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>&lt; .001</td>
<td>0.54</td>
<td>1.62</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.05</td>
<td>-0.73</td>
</tr>
<tr>
<td>4</td>
<td>7.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>263.38</td>
<td>0.77</td>
<td>1.66</td>
<td>-0.81</td>
<td>-0.01</td>
<td>-0.31</td>
<td>-0.05</td>
<td>-0.42</td>
</tr>
<tr>
<td></td>
<td>&lt; .001</td>
<td>0.45</td>
<td>0.12</td>
<td>0.42</td>
<td>-0.01</td>
<td>0.76</td>
<td>0.66</td>
<td>0.29</td>
</tr>
<tr>
<td>5</td>
<td>6.50</td>
<td>0.02</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>-0.06</td>
<td>-0.03</td>
<td>-0.73</td>
</tr>
<tr>
<td></td>
<td>428.36</td>
<td>0.54</td>
<td>0.36</td>
<td>1.53</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.73</td>
</tr>
<tr>
<td></td>
<td>&lt; .001</td>
<td>0.53</td>
<td>1.78</td>
<td>0.13</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.73</td>
</tr>
<tr>
<td>6</td>
<td>6.92</td>
<td>0.03</td>
<td>0.02</td>
<td>0.09</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>247.56</td>
<td>0.08</td>
<td>0.84</td>
<td>4.15</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>&lt; .001</td>
<td>0.94</td>
<td>0.08</td>
<td>&lt; .001</td>
<td>0.02</td>
<td>0.46</td>
<td>0.90</td>
<td>0.47</td>
</tr>
<tr>
<td>7</td>
<td>0.04</td>
<td>0.05</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>-0.88</td>
<td>0.38</td>
<td>-0.88</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.88</td>
<td>0.38</td>
<td>0.38</td>
</tr>
</tbody>
</table>

\[^{31}\]Mixed effects linear regression models used

\begin{verbatim}
lmer(logR1 ~ Voice*WordOrder*Experiment + (Voice+WordOrder+Experiment|Participant) + (Voice+WordOrder+Experiment|ItemNumber))
lmer(logR2 ~ Voice*WordOrder*Experiment + (Voice+WordOrder|Participant) + (Voice+WordOrder|ItemNumber))
lmer(logR3 ~ Voice*WordOrder*Experiment + (Voice+WordOrder|Participant) + (Voice+WordOrder|ItemNumber))
lmer(logR4 ~ Voice*WordOrder*Experiment + (Voice+WordOrder|Participant) + (Voice+WordOrder|ItemNumber))
lmer(logR5 ~ Voice*WordOrder*Experiment + (Voice+WordOrder+Experiment|Participant) + (Voice+WordOrder+Experiment|ItemNumber))
lmer(logR6 ~ Voice*WordOrder*Experiment + (Voice+WordOrder+Experiment|Participant) + (Voice+WordOrder+Experiment|ItemNumber))
lmer(logR7 ~ Voice*WordOrder*Experiment + (Voice+WordOrder+Experiment|Participant) + (Voice+WordOrder+Experiment|ItemNumber))
\end{verbatim}
3.3.4. Discussion

The purpose of Experiment 5 was to investigate whether predictive processes are a core component of the information processing system (i.e., the Prediction-as-Core Hypothesis) by examining whether predictions are generated in real-time, even in cognitively constrained contexts. To simulate a cognitively constrained context, participants were asked to read sentences while retaining a series of letters in working memory. I hypothesized that overall reaction times would be elevated across the board in the task with memory load. In addition, following the Prediction-as-Core Hypothesis, the gradient pattern of sentential patterns would be expected in Region 3, as part of the routine and active generation of expectations by the processing system. Following the findings from Experiment 3, I also predicted to observe the same gradient pattern in Region 5.

Before the hypotheses related to the Prediction-as-Core hypothesis are discussed, I first interpret the effects found in the first two sentence regions. Contrary to the hypothesis, the magnified reaction times were only observed in Region 1 and did not result in interaction effects in any of the other sentence regions. Despite obtaining elevated reaction times only in Region 1, it seems less likely that the impact of memory load was minimal in processing the rest of the sentence regions. The reduction of sharp contrasts between conditions in the later sentence regions suggests that the memory load did have an impact in the later sentence regions, albeit its influence might not have been strong enough to be statistically observed. The magnification of reaction times in Region 1 as compared to other sentence regions was perhaps an artifact of the strategies that the comprehenders consciously employed as they attempted to retain the additional load in working memory while reading the material in Region 1. The significant interaction effect between Voice and Experiment in Region 1 is presumably a spurious effect, since comprehenders had not encountered any argument, verb, or voice-related material at this point in the sentence.

<table>
<thead>
<tr>
<th>Region 7</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.49</td>
<td>0.02</td>
<td>282.62</td>
</tr>
<tr>
<td>Voice</td>
<td>0.01</td>
<td>0.02</td>
<td>0.67</td>
</tr>
<tr>
<td>WordOrder</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.44</td>
</tr>
<tr>
<td>Experiment</td>
<td>0.01</td>
<td>0.03</td>
<td>0.37</td>
</tr>
<tr>
<td>Voice:WordOrder</td>
<td>0.03</td>
<td>0.02</td>
<td>1.33</td>
</tr>
<tr>
<td>Voice:Experiment</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.46</td>
</tr>
<tr>
<td>WordOrder:Experiment</td>
<td>0.01</td>
<td>0.03</td>
<td>0.34</td>
</tr>
<tr>
<td>Voice:WordOrder:Experiment</td>
<td>0.06</td>
<td>0.05</td>
<td>1.29</td>
</tr>
</tbody>
</table>
For both Experiment 3 and Experiment 5, a robust main effect of Voice was observed in Region 2. As the comprehenders encountered the verb with the voice morphology, they consistently demonstrated shorter reaction times for the strongly preferred patient voice. The replication of this effect in Experiment 5 contributes to the strong evidence for the preference of native Tagalog speakers for the patient voice (Cooreman et al., 1984). In addition, the fact that a significant Voice effect was consistently observed in Region 2 for both experiments suggests that comprehenders were attuned to upcoming voice information. This finding establishes the idea that even before comprehenders encounter any argument-related information about the sentence, comprehenders do have Bayesian priors about what the probable voice is of the sentence. This voice effect is observed with verbs that select an animate agent and patient, so future research is needed to confirm whether the same Voice effect would also be observed with verbs that take inanimate arguments.

It was a notable finding that the predicted cline of sentential word order patterns in Region 3 was minimally identifiable in Experiment 5 compared with Experiment 3, in which the pattern was only marginally significant. The statistical findings from the combined experiments showed a significant interaction effect between Voice and Word Order, reflecting the predicted gradience. Furthermore, the same gradience in the expected order of sentence patterns in Region 5 was still observed in both experiments, although the gradience was more sharply contrastive in Experiment 3 than in Experiment 5. A significant interaction effect between Voice and Word Order was also found in Region 5. While this finding is congruent with the Prediction-as-Core Hypothesis, a critical question remains: if the system truly engages in routine generation of predictions, as the hypothesis suggests, why were the findings in Region 3 of Experiment 3 not strongly supported by the mixed effects model?

The inconsistency of predictive effects in Region 3 of Experiment 3 could be a weak effect that only emerges with greater power, such as when the experiments are combined. While this explanation salvages the Prediction-as-Core Hypothesis, it is insufficient to explain the indistinct pattern of results in Region 3 in Experiment 5. Furthermore, the existence of consistent and robust statistical effects in both experiments, specifically the Voice effect in Region 2 and the interaction
between Voice and Word Order effect in Region 5, runs against the argument that sample size was inadequate to observe effects in Region 3.\textsuperscript{32}

I interpret the results in Region 3 as a weak predictive effect. It could be possible that while the system tends to engage in routine predictions, the outcomes of prediction may not always be successful enough for strong effects to be observed. The system might lead itself to infer the Bayesian posteriors of the unfolding linguistic representation (in this case, the sentential word order pattern) based on the Bayesian priors linked with the verb and voice morphology; however, the information accrued might not be adequate to result in a strong prediction that can be sharply observed in behavioral measures. The tendency to have variable prediction strengths can also be corroborated by the finding in Region 5. The strong prediction hypothesis offered in Experiment 3 expected comprehenders to generate predictions following the predicted cline of sentential patterns at Region 3 and to engage in a complete belief-update, resulting in a lack of difference between sentential patterns at Region 5. This strong prediction account was not observed in either experiment. Instead of observing these two hypotheses in these two regions, a weak interaction effect of Voice and Word Order was observed in Region 3, and the same interaction effect was more robustly demonstrated in Region 5. This finding suggests that the success of Bayesian-based pre-activations (including the capacity to perform belief-updating) on a word-per-word basis may not be immediate and can be variable in terms of strength, thereby weakening the strong Prediction-as-Core account. Hence, the pattern of effects demonstrated in Regions 3 and 5 for both experiments highlight variability in prediction strength and the absence of robust, immediate effects of belief-updating in sentence processing. The General Discussion section in Chapter 4 expands on the interpretation of these findings.

The main effect of Word Order observed in Region 4 of Experiment 3 diminished in Experiment 5, perhaps due to the impact of the memory load task with the agent-first advantage observed in Region 4. The induced memory load may have hampered the comprehenders’ attempt to recognize the agent of the event in the sentence. As was discussed in the results of Experiment 4, a cognitive constraint such as memory load impairs real-time parsing operations involving information uptake, prediction, integration, and belief-updating. However, the results of the mixed effects models showed varying effects of memory load at different points in the sentence. In

\textsuperscript{32}This claim is supported by a separate analysis of Experiment 5 alone, which is provided in Appendix F.
particular, the statistical effects of Word Order in Region 4 was affected compared to the interaction effect of Voice and Word Order in Regions 3 and 5. I speculate the possibility that the effects of cognitive load could vary at different levels of representation. For example, higher-level information, such as semantic information (e.g. the integration of an entity’s thematic role to the event representation in Region 4) might be more affected by induced cognitive load compared to lower-level information, such as syntactic information (e.g., prediction of the linear word order in Region 3, or ease of integration of an unfolding phrase to the accumulated syntactic structure in Region 5). This hypothesis needs to be tested in future research.

Looking at the results from the offline measures, the lack of a contrast across conditions in the memory task signifies that the nature of the task was not affected by the variables tested in the experiment. The consistent effect of Word Order in comprehension accuracy in both Experiments 3 and 5 highlights the agent-first advantage in comprehending these sentences, congruent with good-enough parsing (Ferreira, 2003; Ferreira et al., 2002, 2009; Karimi & Ferreira, 2016) and noisy/lossy model approaches to sentence comprehension (Futrell et al., 2020; Futrell & Levy, 2017; Gibson, Bergen, et al., 2013). However, unlike Experiment 3, the Word Order effect between offline comprehension and the real-time parsing in Experiment 5 did not match in terms of statistical significance. Future research should examine whether cognitive constraints affect offline comprehension and real-time parsing differently, with the latter being more sensitive to the additional cognitive load than the former.

Experiment 5 demonstrated the limits of prediction in the cognitively constrained context of a memory load task. Despite finding a significant interaction effect of Voice and Word Order in the critical Regions 3 and 5, the interaction effect was weaker in Region 3 than in Region 5, suggestive of the variability in prediction strength and the absence of robust, immediate effects of belief-updating in sentence processing. These results also illustrate how cognitive constraints such as memory load can impact the real-time processing of voice at different points in the sentence. Finally, the outcomes are also congruent with the claim that successful prediction in the temporal unfolding of the linguistic signal is bounded. These issues are discussed more in detail in the next chapter.
3.4. Conclusion

This chapter looked at three experiments to investigate whether comprehenders can use voice morphology to predict the specific linear word order of an unfolding sentence, and whether predictive effects would be observed in a cognitively constrained context. The following chapter provides a more thorough integration and discussion of all five dissertation experiments.
Chapter 4  
General Discussion

The goal of this dissertation was to expand our understanding on the nature of predictive processes in sentence processing, an influential concept that is described by some scholars to be the core and unifying principle of the human mind (Bar, 2009; Clark, 2013; Friston, 2010; Lupyan & Clark, 2015). Using a series of experiments involving multiple different methods (to be reviewed shortly), I explored whether the sentence parser can utilize informative grammatical information on the verb, called voice morphology, in anticipating and generating expectations for sentential argument order in a verb-initial language, Tagalog. Addressing this question clarifies issues regarding word order preferences in Tagalog, a debated topic in the syntax of the language. Furthermore, this investigation sheds light on the ability of the parser to immediately generate gradient predictions of sentential linear word order patterns. The outcomes of the research inform contemporary accounts of Tagalog word order, as well as models that describe the nature and mechanisms of prediction in sentence processing. Findings from the research produce knowledge that can be beneficial for other related fields such as language pedagogy, speech-language pathology, and artificial intelligence, among others.

Prior to discussing the implications of the study findings to current models of prediction, I first provide a review of the results from the five dissertation experiments.

4.1. Recap of major findings

This section provides a summary of the hypotheses, study results, and interpretation for each experiment.

Because researchers have differed in their analyses of the preferred word order in Tagalog, and empirical evidence has been somewhat limited, especially for less-preferred voices such as the benefactive and instrumental voice, I first determined the relationship between voice morphology and word order preferences in Experiment 1. These word order preferences translate to probabilities, which in turn can be indicative of what comprehenders could use to generate expectations for sentential argument order. Using a sentence continuation task, I tested two
probabilistic and violable constraints, rooted on the constructs of agentivity (agent-first constraint) and pivothood (pivot-second constraint), postulated to drive word order preferences in Tagalog. The results demonstrated word order preferences in the patient, benefactive, and instrument voice that align with the aforementioned constraints; in the agent voice, two patterns were frequently attested, indicating that neither constraint strongly dominates in cases of conflict. The experimental findings do not only demonstrate the strongly preferred patterns in the language, but also highlight substantial variability in Tagalog word order in all four of the examined voices, emphasizing the need to consider gradient combinations of factors in developing accounts of word order in Tagalog.

Importantly, the results in Experiment 1 clarified the roles of agentivity and pivothood with Tagalog voice and word order preferences and established the link between voice morphology and word order. This link suggests that voice can be a strong predictor of word order. Thus, in **Experiment 2**, I used visual world eyetracking to investigate whether Tagalog comprehenders use voice information to pre-activate the upcoming agent and pivot arguments upon hearing the verb. Previous work has reported that Tagalog comprehenders only use verbal semantics and not morphosyntactic information from voice morphology when anticipating arguments (Sauppe, 2016). I build up from this by controlling for argument animacy and testing anticipation of arguments in four different voices. With agentivity and pivothood playing a critical role in Tagalog voice and word order, I predicted increased looks to agents and pivots upon hearing the verb with the voice morphology, with the highest looks to the agent pivot in the agent voice. Surprisingly, comprehenders demonstrated a lack of preferential looks to the agent or pivot as they heard the verb and voice information; in addition, their gaze patterns revealed undifferentiated looks to all four voice conditions throughout the trial. An unexpected animacy effect dominated through the sentence, an observation that suggests comprehenders were generating a wide range of predictions that diminished distinct gaze patterns for each voice condition.

**Experiment 3** addressed whether comprehenders generate probabilistic expectations for syntactic linear word order patterns as they obtain the verb plus voice information. The occurrence of voice and word order patterns in differing frequencies equates to probabilities which comprehenders may use to generate gradient expectations for the unfolding sentence pattern. Specifically, based on Experiment 1 findings, in the patient voice, the highly preferred Agent-Patient order garners the highest probability, vis-à-vis the Patient-Agent order which entails the
lowest probability. In the agent voice, the existence of two equally preferred patterns (Agent-Patient and Patient-Agent orders) provides each with at-chance probability. Hence, I hypothesized that comprehenders would demonstrate surprisal levels reflecting the gradience of these sentential patterns as soon as they encountered the nominal marking of the first noun phrase, to be captured as an interaction effect of Voice and Word Order. Further, as they accrued information from the verb and the first noun phrase, I predicted that comprehenders would engage in a belief-update, such that the upcoming second argument would be the most likely element for them to encounter. The belief-update was predicted to manifest as a lack of difference in surprisal across the word order patterns. The predicted interaction effect of Voice and Word Order was not observed at the hypothesized region, but emerged instead at a later region, i.e., nominal marking of the second noun phrase. These results demonstrate the non-immediacy of prediction in sentence processing, contrary to accounts offered by several scholars (DeLong et al., 2014; Kuperberg & Jaeger, 2016; Kutas et al., 2011; Yan et al., 2017). The initial non-significant but numerical interaction effect at the nominal marking of the first noun phrase can be interpreted as a weak prediction effect, where the cognitive system has not had sufficient pre-activation such that significant effects would be observed. The significant interaction effect observed at the nominal marking at the second noun phrase is most easily seen as an integration effect, reflective of the ease of integrating new information with the accumulated verb and first noun phrase information. Distinct effects of voice, word order, and the interaction between these two factors were also observed at various regions, suggesting temporal processing of voice throughout the sentence. These points are discussed more extensively under the dynamic framework in the subsequent section.

The next two experiments investigated the functioning of predictive processing in cognitively constrained contexts, with the aim of illuminating its ubiquity in the information processing system. In Experiment 4, I tested whether gradient prediction of sentential word order patterns would still be generated, even when time pressure was induced. I hypothesized a speed-accuracy tradeoff between the tasks with time pressure (500 and 750ms SOA via RSVP), and the task that excluded it (i.e., Exp 3 comprehension task). Further, for each task, I predicted that the cognitive system would still generate expectations following the hypothesized gradience of linear word order patterns, to be reflected in the comprehension accuracy of the comprehenders. The outcomes revealed a significant reduction in comprehension accuracy in the RSVP tasks compared to the
comprehension task in Experiment 3, confirming the hypothesized speed-accuracy tradeoff; however, there was no difference between the two RSVP tasks. A significant effect of Word Order was observed in the Experiment 3 comprehension task, illustrating an advantage for agent-first patterns. This advantage, however, weakened in the RSVP tasks, where only an advantage for the highly expected pattern was observed, namely the patient voice Verb-Agent-Patient word order. The reduction in granularity of the cline of sentence patterns in the RSVP tasks was indicative of the limits of prediction to function in contexts with time pressure. That is, in the course of the rapid temporal unfolding of the sentential context, the cognitive pressure reduced the ability of the parser to generate gradient predictions and integrate information from the context, thereby affecting the comprehension of the unfolding material.

It was unclear from Experiment 4 how the addition of the cognitive constraint affected the temporal processing of voice during the incremental processing of the sentence. In Experiment 5, a memory load task was added in the self-paced reading experiment to investigate whether gradient expectations would still be generated in real-time even in cognitively constrained contexts. Specifically, I hypothesized a memory-accuracy trade-off between Experiments 3 and 5 in the comprehension task. Following the claim that predictions are a core principle of the human processing system, I hypothesized that reaction times would demonstrate a gradience in the predicted cline of sentential word order patterns at the nominal marking of the first and second noun phrase regions. This hypothesized gradience would be shown as an interaction effect between Voice and Word Order. The gradience in the nominal marking of the first noun phrase would be presumed to reflect the pre-activation of the unfolding sentential linear word order patterns, while the gradience in the second noun phrase would illustrate the ease of integrating the newly obtained information with the accrued sentential context. In addition, I predicted that there would be a modulation of reaction times for every sentence region as an impact of the induced memory load. Contrary to the first hypothesis, a memory-accuracy tradeoff was not observed, as the overall comprehension accuracy was similar between the two experiments. In an analysis that combined Experiments 3 and 5, the outcomes showed a significant interaction effect between Voice and Word Order for both target regions, with a more robust effect of the predicted gradience in the nominal marking of the second noun phrase than the first. Similar with Experiment 3, I interpret the initial significant interaction as a weak predictive effect of linear word order, while the second
significant interaction is reflective of a strong integrative effect of linear word order. While the modulative effect of memory load to the reaction times was only statistically observed in the first sentence region, the memory load had a sufficient impact to reduce the distinctiveness of the patterns in the critical sentence regions. The results in the final experiment support the non-immediacy of prediction, as well as the variability of prediction strengths in the temporal processing of the sentence.

4.2. The limits of predictive processing

An abundant literature has described prediction as an immediate and fast-acting process. The system is able to use information sources to rapidly anticipate the unfolding signal, as well as to revise beliefs as soon as new information is obtained (e.g., DeLong et al., 2014; Kuperberg & Jaeger, 2016). However, the immediacy of prediction is challenged by the findings in the dissertation experiments. Experiment 2 demonstrated a lack of anticipatory looks to the agent and pivot as two likely arguments of the verb as comprehenders receive verb and voice information. Experiments 3 and 5 illustrated weak predictive effects of predicting the syntactic linear word order of the upcoming sentence as comprehenders encountered verb and voice information, as well as non-immediate effects of performing the predicted belief-update as the verb and first noun phrase information was obtained. Furthermore, Experiment 4 showed how the ability of comprehenders to predict the unfolding sentence was hampered when faced with time pressure; the comprehension accuracy of comprehenders did not reflect the predicted cline of sentential word order patterns. Despite voice being a strong indicator of the likely word order of the sentence, as was demonstrated in Experiment 1, none of the experiments demonstrated successful or strong effects of prediction, suggesting the limits of successful prediction in sentence processing.

The findings illustrating the limits of successful prediction warrant discussion on the factors that engender these limits. This dissertation strongly contrasts with almost all of the extensive research on prediction, which examined predictive effects by providing comprehenders with rich sentential contexts. In fact, in some of these studies, comprehenders encounter at least one or two sentence contexts prior to encountering the target word. Federmeier and Kutas (1999), for example, investigated gradient lexical-semantic prediction and in their study provided comprehenders with rich contexts such as “They wanted to make the hotel look more like a tropical
resort. So, along the driveway, they planted rows of...”. In contrast, participants in the dissertation experiments were only provided part of a verb phrase, i.e., a verb containing the voice morphology (with another adverb or adjunct). This difference brings about three possible factors that have implications on the limits of predictive processing.

The first explanation relates to a time factor, i.e., the match between temporal unfolding of information and the time needed to fully process the information. Comprehenders might need adequate time to process rich verbal information and generate predictions of high granularity. Because different verbs take different voice affixes in Tagalog, it is possible that the interpretation of voice morphology would inevitably tap other verbal information as well, such as subcategorization information. Other research has shown successful prediction of lexical items and lower-level lexical properties under extended time periods, given two arguments, or an argument and a verb (Chow et al., 2018; Chow, Smith, et al., 2016; Huettig & Guerra, 2019; Ito et al., 2016). In the reading experiments (i.e., Experiments 3 to 5), comprehenders might not have been afforded the appropriate time to sufficiently process the verbal information, as they encountered the nominal marking of the first noun phrase immediately after the region containing the verb and voice morphology. It is thus difficult to evaluate whether extra time (such as inserting additional material between the verb and the nominal marking) would have facilitated stronger predictions on the unfolding sentence. However, Experiment 2 has shown that even when comprehenders were provided plenty of time, they still were not able to anticipate the likely two arguments of the verb. In Experiment 2, the verb and adverb region lasted for about 0-1900ms, with an additional extension of 200ms in our analysis for saccadic eye movement. Comprehenders had at least two seconds to process the information and launch eye movements that could be reflective of their likely predictions. While future research can clarify whether successful effects of prediction would be observed using the reading paradigm, the findings in Experiment 2 reduces the likelihood of time being an influential factor in the limited predictive effects observed in the dissertation experiments.

A more appealing explanation for the limited predictive effects relates to the quantity of information available in the sentential context. A certain quantity of information might be necessary for comprehenders to make a successful prediction. The richness of sentential contexts in previous research would definitely contribute to a more successful pre-activation of upcoming
information, as this establishes the higher-level discourse context and/or event representation which could lead to cascading pre-activation at lower levels of representation. As discussed earlier, many earlier studies on prediction provided comprehenders with rich sentential contexts that were one or two sentences long. In some studies, examination of predictive effects involved the use of highly predictable patterns such as *either... or...* (e.g., Staub & Clifton, 2006). In addition, previous studies that examined effects of using verb information to anticipate likely grammatical objects (e.g., Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003), or using verb subcategorization information to predict the upcoming structure (e.g., Arai & Keller, 2013; Boland et al., 1995; Linzen & Jaeger, 2016; Trueswell et al., 1993) had at least one argument (such as the sentential subject) along with the verb to facilitate the generation of predictions. This point is congruent to claims from other scholars that most studies tend to be “prediction-encouraging” (Huettig, 2015; Huettig & Guerra, 2019; Huettig & Mani, 2016), thereby restricting the examination of factors that instigate successful prediction. In the present study, it may be the case that more information, such as an argument in addition to the verb and voice information, would have been necessary for comprehenders to make successful predictions about the upcoming signal. Other than this dissertation, research that has looked on prediction given a single element, such as a verb alone, is scarce (e.g., Garrido Rodriguez et al., 2017, for Tzeltal, a verb-initial language). Though psycholinguistic work on Tagalog is increasing, existing studies either looked at predictive processing where the sentence context only had a verb (Sauppe, 2016) or an argument head and a verb, such as in relativization (Pizarro-Guevara & Wagers, 2020). This dearth of information warrants future research.

Another explanation for the limited effects of prediction could be attributed to the quality of information available in the signal. Different types or sources of information could also contribute to the success of a prediction. As mentioned earlier, other studies had a sentence stimulus that contained at least two critical sentential elements – most of these had at least an argument (more commonly the subject) and a verb (e.g., Altmann & Kamide, 1999; Kamide et al., 2003; Kukona et al., 2011). A fewer number looked at two arguments, such as a subject and an object (e.g., Chow et al., 2016, 2018). What remains unclear is whether a certain type of argument, coupled with the verb, would give rise to successful predictions. The more heavily studied languages in psycholinguistics involve those with restricted word orders where the sentence context provides
the sentential subject and the verb and the comprehender is left to generate an expectation on the unfolding sentential object. In my experiments, comprehenders had stronger integrative effects of the unfolding linear word order when the verb and the first noun phrase (e.g., a core agent or patient) were obtained. In addition to the possibility of having an argument that can be coupled with the verb as contributing to prediction success, further research can explore whether certain types of arguments (e.g., core vs oblique elements) are necessary for successful prediction to occur. In addition, there is a need to investigate in future research whether predictive effects vary depending on the type of syntactic phenomena under examination.

4.3. A dynamic account of prediction

Setting aside the discussion of factors that illuminated the nature of predictive processing, another critical issue involves the ‘ubiquity’ of prediction. Other related terms such as ‘pervasiveness’ and ‘routine’ are often used by scholars to tackle questions regarding the limits or boundedness of predictions in the system (e.g., Huettig, 2015; Huettig & Guerra, 2019; Huettig & Mani, 2016; Luke & Christianson, 2016). These terminologies do not clarify how predictive processing operates in the system. Despite the consonance between these scholars, all of whom propose that predictive processing is bounded (including the dissertation findings), little is known about the regularity of successful predictive processing in the system and the mechanism that captures its bounded properties. Is the success of pre-activating upcoming material a product of prediction being an all-or-none, deterministic process, or as a gradient activity in the system?

A more cohesive explanation would be to assume that prediction, just like other cognitive operations, is gradient in nature, where its levels of activity are variable, and where its outcomes of successful and unsuccessful prediction also vary depending on the amount of pre-activation the system has accrued. This account of prediction is compatible with a description of a dynamic cognitive system. In a dynamic framework, the cognitive-linguistic system is construed as a complex system, where elements in multidimensional levels of representation are engaged in constant self-organization. The cognitive-linguistic behaviors that emerge at different points in time are a product of the self-organizing tendencies of the system. Both internal (e.g., grammatization tendencies or constraints of language, processing tendencies such as priming and recency effects found in language) and external factors (e.g., statistical properties of the input)
shape the statistical properties of the cognitive-linguistic system. The activation states of these elements at these multiple levels of representation constantly change over time, resulting in the system’s relative stability or instability. Elements within a level and across levels are highly coordinated, in such a way that activation changes in an element can have cascading effects on the activations of other elements. The dynamic changes that happen at one point in time are linked to the changes that will happen at later time points (Holt & Osman, 2017; L. B. Smith & Thelen, 2003).

Working against this backdrop of a dynamic system, predictions for bottom-up linguistic structures can be construed as dynamically changing pre-activations over time. These pre-activations are graded; the different sources of information adapted by the system are construed as cascading activations that influence the gradient levels of these pre-activations. These dynamically changing pre-activations encompass multiple levels of representations. This view contrasts with descriptions of prediction as a discrete, all-or-none pre-activation of the upcoming lexical item (e.g., DeLong et al., 2014; Van Petten & Luka, 2012).

Furthermore, while the system can be engaged in extensive amounts of pre-activation over time, the strength or impact of a prediction to the developing syntactic analysis or linguistic representation can be characterized as having tipping points. The activation states of certain representations require a tipping point – that is, a certain level of activation across multiple levels in the system – to make a successful or strong prediction. When a tipping point is met, the parser is ready to perform the subsequent analyses for that developing representation. Hence, the success of a prediction can be captured when the dynamically changing activation levels matches the tipping point.

A dynamic system captures the variability of pre-activation observed across multiple levels of representation. For instance, in Experiments 3 to 5, besides pre-activating the likely syntactic representation in the nominal marking of the first noun phrase, it can also be surmised that on that same region, there can also be a possible pre-activation of the upcoming lexical item, as well as

---

33The term ‘tipping point’ is distinguished from ‘threshold.’ The latter often tends to be used in the psycholinguistic literature to describe the minimum activation within a single level of representation that is needed to initiate a cognitive process. ‘Tipping points,’ as used in this context, describe activations that encompass multiple levels of representation. When sufficient activation is achieved, a tipping point is met, which leads to the initiation of a subsequent cognitive process.
the pre-activation of the phonological features of the upcoming lexical item. A pre-activation at the higher levels, such as anticipated thematic roles or event structure, can also be assumed. As reviewed in the Background, successful prediction at different levels have been reported in two decades of research (e.g., syntactic: Lau et al., 2006; Staub & Clifton, 2006; Van Berkum et al., 2005; Wicha et al., 2003; Yoshida et al., 2013, lexico-semantic: Federmeier & Kutas, 1999; Kutas & Hillyard, 1984; Staub et al., 2015, and phonological/orthographical: DeLong et al., 2005; Kim & Lai, 2012; Laszlo & Federmeier, 2009). The variability by which tipping points are met across multiple levels suggests that the granularity and rates of successful pre-activations across various levels differ. The probability of successful pre-activation at higher levels of representation (e.g., the likely semantic and syntactic representation of the unfolding event) could be higher compared to successful pre-activation at the lower levels (e.g., the pre-activation of the lexical item or the specific phonological features of the segments of the lexical item, particularly in a context that includes a visual scene, or other highly constraining material). This difference could be attributed to the number of alternatives available in the probability space, as well as the weights associated by the Bayesian priors with these alternatives. The probability space of the alternative candidates for the semantic and syntactic representations could be smaller compared to the alternatives for the lexical and phonological features.

It is also important to note that scholars have used the term ‘dynamic’ in different senses. For example, Mani and colleagues (2016) utilize the term ‘dynamic’ to refer to the employment of multiple systems in prediction (e.g., production, association, combinatorial systems, and event simulations; Huettig, 2015). The dynamic prediction account discussed in this dissertation is more congruent with Kuperberg and Jaeger's (2016) and Kuperberg's (2016) definition of prediction as the use of top-down information to generate a higher-level representation and pre-activate the upcoming bottom-up input. Despite the brief mention of the term ‘dynamic’ in these two latter articles, the dynamic account discussed here is sympathetic to theirs, and extends it further by illustrating how the dynamic system explains successful and unsuccessful predictions.

4.4. Complementing prediction with integration in a dynamic account

As has been mentioned in sections 1.1.4.4 and 4.2, the sentence processing literature has been dominated by work on predictive processing. While predictive processing is assumed to play a
major role in comprehension, its complementary process, integration, should be equally acknowledged as an influential factor in sentence processing. Integrative processes are broadly defined as the ease of assimilating a newly obtained structure to the accrued information or the developing representation in working memory. In classic research, integrative processing has been heavily investigated with syntactic ambiguity resolution, i.e., the ease of incorporating a new structure in the tentatively developing syntactic analysis. These integration-based perspectives were then predominantly incorporated in serial, garden-path models (Frazier, 1979; Frazier & Clifton, 1996; Frazier & Clifton Jr, 1989; Frazier & Rayner, 1982), parallel, lexicalist models (MacDonald et al., 1994; Seidenberg & MacDonald, 1999; Trueswell et al., 1993), and hybrid models that incorporate components from both serial and parallel models (Boland, 1997; Boland et al., 1995; van Gompel et al., 2001).

As has been demonstrated in my dissertation experiments, not all outcomes of predictive processing are successful; integration processes have also been found to play a critical role in the comprehension process. Thus, it is logical to incorporate integrative processing in the dynamic account as well. Integration has been useful in situations where the system requires further processing of accrued information prior to the implementation of the subsequent operation. For example, in Experiment 3, a strong prediction account predicts the generation of an expectation for the upcoming syntactic linear word order in the nominal marking on the first noun phrase, and hypothesizes a belief-update once the nominal marking of the second noun phrase is encountered. Observations in these two regions were not accounted for by the strong prediction account. While a weak predictive effect was observed in the nominal marking of the first noun phrase, there was a consistent robust integrative effect observed in the nominal marking of the second noun phrase region. These two effects were attested in Experiment 3, and were replicated in Experiment 5.

Integrative processes have also shown to be manifest in cases where the newly acquired information enriches the developing syntactic analysis or event representation. The findings in the subset analysis of Experiment 2 extend support for this claim. Recall that anticipatory effects of the agent and pivot as the two likely arguments of the verb were negligible as comprehenders obtained the verb and voice information, and the comprehenders’ variability of predictions for the unfolding sentence was speculated as the possible cause for this finding. In the subset analysis involving the patient voice, as comprehenders obtained additional information from the first noun
phrase (i.e., the agent as the first argument), they increased looks to the patient pivot as it was being mentioned as the second argument. This pattern of increased looks to the Patient demonstrates the confirmation of the Patient pivot as the second argument. Looks to the patient decreased in the subsequent regions, suggesting an update in comprehenders’ beliefs that the patient was not likely to be mentioned in the subsequent regions. This reduction in looks to the patient after the second noun phrase reflects an adjustment to the range of predictions participants were making regarding the unfolding signal.

While some recent work acknowledges the complementary role of integration in comprehension (Ferreira & Chantavarin, 2018), prediction has dominated recent sentence processing literature, reducing attention to the role of integration in the comprehension process. Under a dynamic account, processes such as prediction and integration supplement and resemble each other. Both operations can use multiple sources of information to impact the activation for the developing representation, and both can encompass multiple levels of representation. Both can use the generated higher-level representation to dynamically change the lower-level activations in the anticipation of upcoming material (for prediction), or in the incorporation of newly received input (for integration). The only difference between the two is the timing by which input is received. As discussed in Chapter 1, prediction or pre-activation refers to anticipation prior to receiving the signal, while integration is the effect observed once the input has been received. Under this dynamic framework, the side-by-side implementation of prediction and integration enriches the higher-level representation in the temporal processing of the unfolding signal.

A dynamic account is advantageous for two reasons. Notably, a dynamic account avoids confusion in terminological use. For instance, some scholars define anticipation as reflecting the gradient activation of lower-level features, while prediction refers to the deterministic, all-or-none identification of the upcoming specific lexical item (DeLong et al., 2014; Van Petten & Luka, 2012). However, there is reason to believe that the activation of representations across levels are gradient during processing, eliminating the need to separate anticipation from prediction. As reviewed in the introductory chapter, the prediction of a specific lexical item is best described as gradient rather than deterministic (Luke & Christianson, 2016; Staub et al., 2015). Previous research has also described gradient activation of phonological/orthographical and lexical-semantic features of the upcoming word (DeLong et al., 2005; Federmeier & Kutas, 1999b; A.
Kim & Lai, 2012; Kutas & Hillyard, 1984; Laszlo & Federmeier, 2009; Paczynski & Kuperberg, 2011; Staub et al., 2015; Wang et al., 2020). This dissertation contributes evidence to the gradience of syntactic representations, with the predicted cline of sentential linear word order patterns in Tagalog as a case of gradience. Thus, the extensive evidence on gradience that occurs at various levels of representation supports the dynamic view, which claims gradience in predicting across all linguistic levels.

A dynamic account also proposes that the processing operations are gradient as well. The findings from more recent studies highlighting the limits of prediction (e.g., Chow, Momma, et al., 2016; Huettig & Guerra, 2019; Ito et al., 2018), including the current dissertation findings, support the claim that prediction outcomes can be successful or unsuccessful. The success of prediction outcomes, which can be described in a continuum of strength, is best described quantitatively. In Bayesian terms, predictive processes could be computationally represented as the strength of the prior distribution to probabilistically infer the posterior distribution of the unfolding signal. A strong predictive effect results in sufficient activation to reach the tipping point of the predicted representation (Kuperberg, 2016; Kuperberg & Jaeger, 2016; Yan et al., 2017). Integrative processes could be viewed under the same Bayesian computational formalism as well, i.e., the strength to shift the probabilistic distribution of the posterior after the information has been received. A strong integrative effect obtains adequate activation to reach the tipping point of the consolidated representation. Hence, a dynamic account unifies the mechanics of the human information processing system, by viewing the whole system – both the linguistic representation and the processing system – as operating in a gradient fashion.

4.5. Illustration of the dynamic account using the dissertation experiments

The dynamic prediction mechanism can be illustrated with the outcomes from Experiments 2 to 5. (Note that for the purposes of simplicity, I frame these representations in this example as categorical, although they are gradient.) Experiment 2 aimed to examine whether comprehenders immediately use voice to anticipate the two likely arguments of the verb, i.e., the agent and the pivot. Viewed through a prediction perspective, the goal was to observe whether comprehenders could use the voice morphology on the lexical verb to generate a higher-level event representation, where the agent and pivot elements would be pre-activated roles in the event. The results showed
that comprehenders did not demonstrate immediate anticipatory effects of voice; only animacy effects were observed. This implies that there could be very minimal pre-activations to the agent or the pivot element relative to the other elements at the time that comprehenders obtain the verb information. The tipping points for the gaze patterns to anticipate have not been met, resulting in an unsuccessful prediction. The speculation that comprehenders generate multiple predictions in parallel imply that multiple representations are activated, resulting in an unreached tipping point for a single representation (i.e., an agent-first, pivot-second word order).

In the same experiment, there was also evidence of variability in the accumulation of activation for the likely unfolding arguments, as a function of the strength of the cues available in the signal. After the first noun phrase is encountered, the system has obtained information from the verb containing the voice morphology, as well as the agent being the first argument. This information increases the boost for the activations for the pivot as one of the verbal arguments, as well as the next argument in the linear word order of elements. However, the frequency difference between patient and benefactive voice renders differences in terms of the strength to activate the patient/benefactive as the second sentential argument. The highly frequent patient voice obtains an accumulation of activation for the patient pivot and attains the tipping point for the salience of the pivot, which translates to significantly increased looks to the patient pivot as it was mentioned in the signal. This was observed in our results as a significant interaction between Voice and Region in the subset analysis of the NP2/NP3 regions. However, in the case of the benefactive voice, accumulation of activation for the benefactive is less than that of the patient pivot, since the benefactive voice is a less frequent pattern compared to patient voice. Hence, the generation of inferences for the posterior probability for the benefactive would entail less reliability for this pattern compared to the highly frequent patient voice. Hence, the pattern of increased looks to the benefactive pivot would be less compared to the patient pivot. Activation to the benefactive may not be sufficient to reach the tipping point; this difference is captured as a non-significant interaction effect of Voice and Region for BV, in contrast to the significant interaction effect of the two factors for PV.

Experiment 3 tested whether comprehenders could generate an expectation for likely sentential word order. There could be two possible kinds of predictions. A coarse-grained prediction would result in an anticipation of the patient voice of the Verb-Agent-Patient word order, since this
pattern is the most frequent in the language. In contrast, the hypotheses offered in Experiment 3 reflect fine-grained predictions, where the incorporation of voice information results in an inference of the unfolding signal that mirrors the predicted cline of sentential word order patterns. This goal can be mechanistically represented as the pre-activation of a more likely, versus less likely, higher-level syntactic representation, using voice information. The results in the nominal marking of the first noun phrase of Experiment 3 demonstrated numerical differences but a lack of a significant interaction effect of prediction, suggesting possible pre-activation of the likely word order patterns, but at an unreliable level across trials. This resulted in a generally unsuccessful or weak prediction at this point. That is, as the system obtained verb and voice information, it could have begun to shift to activate linear word order patterns following the predicted gradience; however, the accumulated pre-activations may not have been strong enough to result in robust differences in surprisal values, or to implement a belief-update prior to when the nominal marking of the second noun phrase is reached. The later interaction effect of voice and word order, found at the nominal marking of the second noun phrase, suggests that the accrual of additional information from the first noun phrase accumulates activation for these gradient linear word order patterns. A tipping point was reached as this second nominal marking was observed: the accumulated activation of the different linear word order patterns was adequate to shift the probabilities and anticipate the unfolding nominal in the subsequent region as the highest likely element. This belief update is observed as a lack of difference between linear word order patterns in the next region. This accumulation of activation, however, is not a reflection of pre-activation, since information about the word order has already been obtained from the first noun phrase, and more than the expected input has already been received by the time the effect is observed. Note that these same patterns of weak prediction and strong integration of linear word order in Experiment 3 were observed with the findings in Experiment 5, with the only difference being that Experiment 5 was implemented with memory load.

Experiment 4, which tested whether comprehenders could generate expectations about the sentential linear word order when faced with time pressure, showed comprehension accuracies that did not match the predicted cline of sentential word order patterns. As discussed in the Discussion section of Experiment 4, comprehenders are expected to perform information uptake, to generate predictions, to integrate the input, and to perform a belief-update as they process the sentence. As
comprehenders receive the verb and voice information, they could pre-activate representations of the likely sentential word order patterns based on their hypothesized cline. Instead, the results indicated that the system cannot cyclically perform fine-grained expectations, integrations, and belief-updates on the unfolding signal given the speed by which the sentence stimuli are presented. The system resorts to a tendency to generate a coarser-grained expectation for the post-verbal regions which is not sensitive to voice information. It thus generates a prediction for the patient voice of the Verb-Agent-Patient word order. It is likely that the activations for the representation of the coarse-grained expectation remain intact and are not updated as more information is accrued. No tipping point is reached, and the anticipated form is perhaps reinforced as the input unfolds. Another possibility for this situation would be for the comprehender to still have generated fine-grained expectations for the unfolding signal, but that the incorrect accuracy response reflects possible problems rooted in later memory processes, such as experiencing a decay of the activated linear word order representation.

4.6. Online processes and offline comprehension

Another relevant matter to discuss is the match between real-time comprehension, measured by gaze patterns and reaction times, and offline comprehension, measured by responses to comprehension questions. The dissertation experiments showed mismatches between real-time processing and offline comprehension, in the sense that, given the responses to the comprehension question, the details of processing the sentence stimuli do not always reflect the most likely representations. These mismatches have been documented in previous research, where comprehenders retained interpretations of incorrect syntactic parses (as observed in comprehension responses) despite evidence for syntactic reanalyses observed from reaction time data (e.g., Christianson et al., 2001, 2006; Patson et al., 2009). These mismatches are critical in understanding the bigger picture of the comprehension process (Ferreira & Yang, 2019). Before I proceed to the implications of these mismatches for the dynamic account discussed in this chapter, a recap of the mismatches between real-time processing and offline comprehension is provided.

In Experiment 2, the lack of agent and pivot effects upon encountering the verb plus voice morphology, as well as the lack of differentiated gaze patterns across voices, may be suggestive that comprehension across these patterns would be similar across voices. However, we observe
better performance in terms of comprehension accuracy between agent and patient voice than benefactive and instrumental voice. This difference between voices could be attributed to a frequency difference, i.e., agent and patient voices are more frequent than the benefactive and instrumental voices (Himmelmann, 1987). This potential difference between the more and less frequent voice patterns is somewhat captured in the gaze patterns in the subset analysis, by the greater looks to the mentioned elements in the more frequent voice (PV) compared to the less frequent voice (BV). Despite this finding, the gaze patterns are not sharp enough to illustrate contrasts between voices, making the link between real-time processing and offline comprehension unclear.

The reaction time measures in Experiments 3 and 5, reflecting various predictive and integrative effects, demonstrated the real-time processing of voice throughout the sentence. Despite the integration of all argument information at the end of the sentence from real-time processing measures, the comprehension accuracy responses showed a Word Order asymmetry, i.e, a greater advantage for Verb-Agent-Patient patterns than the opposite word order. Meanwhile, for the RSVP experiment in Experiment 4, comprehenders neither showed this agent-first word order advantage nor the predicted cline of sentential patterns in their comprehension accuracy in the conditions involving time pressure. Instead, they were observed to have the highest accuracy in the most frequent pattern, which is patient voice of the Verb-Agent-Patient pattern. Unlike in Experiment 2, the comprehension questions in Experiments 3 to 5 evaluated comprehenders’ comprehension of the material they read by testing their knowledge of the thematic roles of the event entities (i.e., the who-did-what of the event; see Methods Section 3.1.2.3 regarding the construction of the comprehension questions). This pattern is consistent with a dynamic processing system that does not always result in a fully veridical parse.

In the discussion sections of the former chapter, these offline comprehension effects in Experiments 3 to 5 were explained using two accounts: the “good-enough” approach and the noisy-lossy-context account. The “good-enough” approach proposes that comprehenders apply heuristic strategies in their comprehension of the sentence, thereby producing a superficial processing of the sentence material (Christianson et al., 2001, 2006; Ferreira, 2003; Ferreira et al., 2002, 2009; Patson et al., 2009). In a more recent account, Karimi and Ferreira (2016) introduced the notion of “cognitive equilibrium” in their “good-enough” account, suggesting that comprehenders
simultaneously process a sentence using two routes: a heuristic route, where heuristics are applied in comprehending the material; and, an algorithmic route, where the bottom-up processing of the unfolding input is analyzed using grammatical knowledge. The system attains “cognitive equilibrium” after the heuristic route and an interim output is generated; however, if the equilibrium becomes unstable because of the strong evidence from the algorithmic processing route, the interim output is further revised. Following this proposal, the asymmetry between the Word Order effect in comprehension accuracy and the wrap-up effects observed at the end of the sentence in real-time processing in Experiments 3 and 5 may suggest the tendency for heuristic analysis (i.e., an agent-first analysis) to dominate the interim output. The stability of the interim output from the heuristic route may be so strong that the need for reanalysis from the algorithmic route does not influence the interim output. In Experiment 2, the mismatch between the more and less frequent voice patterns may also reflect a dominating heuristic analysis in the interim output, i.e., the frequent Verb-Agent-Patient analysis (applicable to patterns used for both AV and PV in the test sentences), with the output being so stable that a reanalysis is not achieved. In both of these cases, the “above chance” performance that matches with the heuristic analysis was observed for those patterns.

In the case of Experiment 4, the immediate heuristic analysis (i.e., the most frequent voice and word order pattern) dominates in conditions with time pressure because of weak evidence obtained from algorithmic analysis. Because the system is unable to keep up with the processes of information uptake, prediction, integration, and belief-update, the analysis from the algorithmic route is less reliable, and the interim output is dominated by the analysis from the heuristic route.

From a noisy-/lossy-context standpoint, noise distorts the processing, thereby affecting comprehension of the material. These two accounts differ on where the noise is situated, which has implications on the computational formalisms for these models. In noisy-context surprisal, the noise affecting the processing is situated in the sentence context (Bergen et al., 2012; Futrell & Levy, 2017; Gibson et al., 2013; Gibson et al., 2013, 2017; Hahn et al., 2019). From the noisy-context perspective, it could be possible that comprehenders may have mistaken the nominal marking encountered on the nouns (e.g., switching ang [ʔaŋ] for ng [naŋ]), thereby resulting to an incorrect parse. A possible cause of the misinterpretation might be to employ a coarse-grained prediction, such as to automatically interpret the first noun as the agent, or to interpret the first
nominal marking with *ng*, the marker for the agent in PV. In contrast, the lossy-context perspective incorporates a noise component in the memory representation of the sentential context (Futrell et al., 2020). As an example, an imperfect parse as a result of the misinterpretation of nominal marking results in a generation of an incorrect event representation, where the nominal entities are assigned the incorrect thematic role. The error could also stem from a correct parse of the input, but an application of a heuristic that produces an incorrect higher-level memory representation.

All of the experimental findings can also be explained with the noisy-/lossy-context account. In Experiment 2, the asymmetries between the more frequent and less frequent voices can be attributed to a prediction that utilizes the coarse-grained prediction (e.g., agent-first word order, or an agent acting upon a patient as a typical semantic interpretation), producing an incorrect parse or an incorrect memory representation. A similar superimposition of a coarse-grained prediction (e.g., agent-first word order) to the developing parse or higher-level representation can also be construed for Experiment 3, which causes errors in comprehension. Furthermore, noise from the nature of the task can also result to noisy-/lossy-context processing, as has been illustrated in Experiments 4 and 5. The rapid presentation of sentence stimuli in Experiment 4 mimics the noise in the sentence input when doing speed reading. Similarly, the performance of a simultaneous memory load task while reading sentence stimuli in Experiment 5 simulates divided attention when multi-tasking. Thus, the induced cognitive load can be likened to noise that can affect the accuracy of the parse and the developing higher-level representation as the sentence is processed.

The dynamic account also offers an explanation for the mismatch between real-time and offline comprehension results. The dynamic account supports the “good-enough” and noisy-/lossy-context surprisal accounts, and further extends these accounts by providing a more detailed explanation of the mechanisms that occur at the algorithmic level. A mismatch between real-time and offline comprehension can be explained by an incorrect prediction that continues to linger throughout the sentence. I will provide examples of these cases from the Experiment 3 and 5 results. Suppose a comprehender encounters patient voice morphology on the verb, and generates two representations of the linear word order in parallel, with one representation being much more likely (i.e., Verb-Agent-Patient) than the other (i.e., Verb-Patient-Agent). At the same time, the comprehender generates a higher-level semantic and event representation (e.g., an agent acts upon the patient) that is in consonance with the activated syntactic representation (i.e, highly probable
Verb-Agent-Patient) and where the first nominal is parsed as the agent. As the signal containing the nominal marking of the first argument unfolds (e.g., patient marking ang), the comprehender’s gradient predictions reflect a weak surprisal difference between the two representations, and the difference strengthens until she obtains confirmatory evidence from the nominal marking at the second noun phrase (i.e., that the pattern from the input is the unexpected Verb-Patient-Agent). As the comprehender ends the sentence with a wrap-up of all argument information, it might be the case that the incorrect semantic and event representation is not updated. Thus, despite the integration and belief-update observed at the syntactic level, the higher-level representations remain unchanged, thereby resulting in an error in interpretation. Hence, the Word Order asymmetry in Experiments 3 and 5 may be reflective of the comprehender’s tendency to adapt the initial higher-level semantic and event representation parse (i.e., agent-first, patient-second alignment), without being capable of a belief-update, especially in sentences containing patterns that are incongruent with the initial prediction.

Another example of an incorrect parse can be illustrated with Experiment 4. Comprehenders could generate fine-grained (albeit weak) predictions of the unfolding sentential linear word order as they obtained the verb and voice information, but due to the speed by which the argument information was presented, the system was unable to hold information in working memory and perform the necessary operations, such as pre-activating subsequent input, integrating accrued material, and performing a belief-update. The comprehenders merely adopted the coarse-grained prediction for the higher-level representation until the end of the sentence, and the processing system was unable to perform a belief-update because of the unreliable priors from the syntactic parse.

In the same experiments, errors may have also been brought about by a correct parse of the material, and an incorrect parse of the question, or effects of later memory processes. In this context, the syntactic and higher-level representations generated by the comprehender during pre-activation are in synchrony, and they continue to be congruent as the sentence wraps up. However, an incorrect parse of the question may lead the comprehender to assume that it matches with the parse of the sentence material, thereby producing an incorrect comprehension response. Or, the activated syntactic representations may have decayed, which makes the comprehender unable to access the representation that is faithful with the linguistic stimulus.
As shown in these examples, a dynamic account of comprehension complements “good-enough” and noisy-/lossy-context models as it provides an algorithmic explanation of the trajectory of developments that occur at multiple levels of representation during real-time processing. Whether the trajectory of activations change across levels can have an impact on comprehension success. I believe that the dynamic account has the potential to offer other explanations beyond what is offered here in this sub-section.

4.7. Limitations and future research suggestions

In order to further expand our knowledge on the word order phenomenon in Tagalog, as well as the role of predictive processing in sentence comprehension, several aspects warrant further investigation. I recommend the following to be explored in future research.

Further examination of Tagalog word order preferences. Experiment 1 examined the word order preferences in four different voices (agent voice, patient voice, benefactive voice, and instrument voice), and two probabilistic constraints that explain the dominant word order preferences in the language. This study extended work on word order preferences by investigating how these constraints apply to understudied patterns, such as the benefactive and instrumental voice. Because the emphasis of the research question has been on transitive verbs, it would be relevant to examine other patterns as well, such as the locative voice, and other verb types, such as psychological verbs. It would also be useful to explore the consistency of the experimental findings with oral and written corpora, though these data sources are still in the process of being developed for Tagalog.

The current study has focused on agent saliency and pivothood as two primary factors influencing word order; however, as summarized in the introductory chapter, several other factors can also influence word order. Factors such as the thematic hierarchy in Philippine-type voice systems, and other syntactic constraints such as the core/oblique distinction also warrant further examination. Because the current experiment controlled for argument animacy, it would be interesting to explore whether similar results would be observed when the animacy of the arguments are manipulated (e.g., inanimate agent, animate patient). Future research could also investigate whether parallels of these word order tendencies exist across discourse contexts or in
natural discourse, or how information structure and prosodic length play a role and interact with these grammatical constraints on word order preferences.

One of the salient results in the experimental findings is the disparity between AV/PV and BV/IV. In particular, there was a higher number of production errors in BV/IV than AV/PV. A further analysis of scores showed that the large number of errors in the benefactive and instrument voice patterns were almost evenly distributed across participants and items. While the higher error rates in the benefactive and instrumental voices were attributed to potential factors such as their relatively lower frequency (Himmelmann, 1987), higher processing difficulty, or need for a more constrained or felicitous discourse context, additional research is necessary to provide support for these claims.

Other scholars have also observed a large asymmetry between agent and non-agent voice patterns. In non-agent voices, both the agent-first and pivot-second tendencies can influence the linear word order simultaneously, while in the agent voice only one of these tendencies can be instantiated at a time. Scholars have offered explanations for the distinction between agent and non-agent voice patterns (Aldridge, 2016; Kaufman, 2009, 2017; Ross, 2009), and the theoretical implications of this asymmetry could be explored in future research. In addition, there was a weak tendency to mention the patients in the early positions in the BV and IV. The role of the “patient-early” preference, in comparison to the other constraints, could be probed as well in future research studies.

Another interesting aspect of the results is the contrast observed in linear word order preferences between native Tagalog-speaking adults versus children (Bautista, 1983; Garcia, Dery, et al., 2018; Segalowitz & Galang, 1978; N. Tanaka, 2016) and versus adults with agrammatic aphasia (Bondoc et al., 2018). In the latter two populations, an overall Agent-Patient preference was observed in production. It may be the case that these populations attune their cognitive resources, or implement psycholinguistic processing, with mechanisms different from those of native Tagalog adults. However, the set of studies differ in several aspects of design, including sample sizes for participants and items, and type of stimuli (e.g., argument animacy differences). As this is not the primary focus of the dissertation, further investigation is warranted.

The role of constraints across processing phenomena. The constructs of agent saliency and pivothood have been key to understanding the dominant word order patterns in Tagalog. The
influences of agent saliency and pivothood (i.e., syntactic prominence) have long been established in cognitive and theoretical linguistic research. However, much remains unknown about how these factors work in sentence processing, especially when it comes to languages with understudied typological properties. More recent work has explored how pivothood plays a role in processing filler-gap dependencies (Pizarro-Guevara & Wagers, 2020). I recommend the examination of agentivity and pivothood with regard to processing other syntactic phenomena. Future research could also investigate whether the relative strength of animacy versus agent prominence and pivothood would vary depending on the syntactic phenomena and discourse context.

Influences to successful predictive processing. A major finding from the dissertation experiments is the limits of successful prediction in sentence processing, which were attributed to the amount and type of information available in the signal to make successful predictions. There is still a need to further examine how these factors could influence the success of a prediction. Below, I elaborate on possible avenues for the future exploration of these factors.

The lack of agentivity and pivothood effects using verb and voice information in Experiment 2 was attributed to the wide range of predictions the comprehenders were making. In future studies, the grammatical role (i.e., argument vs adjunct), number, and animacy of the entities could be controlled in order to further probe the possible effects of agentivity and pivothood in predicting the upcoming arguments. The examination of these effects could be explored across verbs of varying transitivity and voice. Exploring the separate effects of voice and verb information in predicting the unfolding arguments is also recommended.

As mentioned in Chapters 3 and 4, the current dissertation varies from the extensive work on prediction in that the experiments only provided comprehenders a limited amount of information (i.e., the verb and voice information) to generate gradient predictive effects of voice. As weak predictive effects of voice were observed in Experiments 3 and 5, it would be worthwhile to examine whether the addition of certain types of arguments (e.g., agent or core argument) besides the verb would increase the likelihood of successful predictions on the upcoming arguments. The examination of various types of syntactic phenomena where the voice and the additional arguments occur (e.g., agent or pivot) may allow the examination of these predictive effects.

While it has been acknowledged that prediction happens at various levels of representation and at different rates, the situations and speed at which successful prediction could happen at these
various levels remains unknown. A comparative analysis of factors that engender successful prediction at different linguistic levels would be relevant in understanding the function of prediction in the cognitive-linguistic system. It would also be useful to investigate the interaction between prediction and integration in the successful prediction of the unfolding signal, as well as in the successful comprehension of the material.

Other sentence processing effects. The word order effect which was observed in Experiment 3 and diminished in Experiment 5 warrants further investigation. It is unclear whether this effect could be solely attributed to an agent-first word order advantage, or to some other factor such as lower-level lexical agent effects. Further studies could clarify how such effects relate to or contrast from syntactic predictions (e.g., prediction of a specific linear word order).

Moreover, as contemporary sentence processing models incorporate the notion of noise and memory effects to the generation of expectations to the unfolding signal, it would also be beneficial to evaluate how the predictive effects observed in this study would differ when the noise factor is incorporated. The current measures used in the reading studies employ the construct of surprisal from expectation-based models in determining predictive effects (Hale, 2001; Levy, 2008), but these effects can be modified once noisy and memory effects are incorporated in the computation of surprisal (Futrell et al., 2020; Futrell & Levy, 2017; Gibson, Bergen, et al., 2013).

The addition of cognitive constraints such as time pressure and memory load has shown how the reduction of cognitive resources can impact sentence processing to varying degrees. Experiments 4 and 5 demonstrated how the prediction of a sentence’s linear word order can be affected by the induced cognitive constraint, such as the reduction in prediction gradience, or the loss of effects at different sentential points as voice is processed. The manipulation of the degrees of these cognitive constraints (e.g., varying SOAs or varying amounts of memory load) in the processing of linguistic material could expand our understanding of the relationships between the different cognitive operations and resources.

Finally, future research could explore the use of other experimental paradigms, such as neuropsychological methods and eyetracking while reading to gain more evidence about the effects observed in the dissertation experiments. Experimental paradigms often provide varying natural language contexts from which language use is examined; hence, the use of various methods
could help corroborate the dissertation experiment results and extend insights about the mechanics of the cognitive system.

4.8. Summary

This chapter provided a recap of the major findings from each dissertation experiment. A common theme across the results is the variability of successful predictive processing. The dissertation findings are viewed using a dynamic account, where linguistic representations are described in terms of trajectories of accumulating activations as temporally shaped by both prediction and integration. The dynamic account also explains performance mismatches between real-time and offline processing, highlighting the contribution of the dynamic account to explain various comprehension processes. The next chapter concludes the dissertation.
Chapter 5

Conclusion

This dissertation extended our understanding of predictive processing, argued to be the unifying principle of the human information processing system. Despite the extensive work on predictive processing, little has been known so far about the ability of the parser to use informative grammatical information to generate gradient expectations beyond the upcoming word, such as the unfolding linear word order of a sentence. This dissertation examined whether comprehenders could use grammatical information on the verb, called voice morphology, to predict the unfolding linear word order of the sentence. (Readers who would like a detailed summary of the dissertation experiments are directed to the first section of Chapter 4.) The findings of the dissertation can be summarized in three main points.

Tagalog word order is gradient. The findings in Experiment 1 highlighted the applicability of the proposed agent-first and pivot-second constraints in driving word order preferences in Tagalog. These probabilistic constraints permit the dominant patterns to be captured in the language. In the non-agent voice patterns, there is a tendency for agents to be in the first position, and the pivots to be in the second position. In the agent voice where the agent is also the pivot, the equal weight of these two constraints permit two similarly preferred patterns in this voice, namely the agent-patient and patient-agent word order patterns.

In addition, the results from Experiment 1 also illustrated smaller counts of mention to other elements in the first and second positions, emphasizing the variability of word order in the language. For example, in non-dominant patterns, the patient element was observed to be mentioned in the first and second positions. Other non-agent and non-pivot elements had smaller counts in these positions as well. These smaller counts produce the non-dominant yet grammatically acceptable patterns in the language.

The dominance and variability of word order patterns demonstrate a gradience of word order patterns in the language. Such gradience captures both dominance and variability and should be accounted for in future descriptive or theoretical work on Tagalog syntax.

The gradience of predictive processing can be observed with the variability of prediction success. The four experiments that examined whether voice morphology could be used
predictively by native Tagalog speakers all showed that predictive processing has its limits, contrary to major claims in the literature that prediction is immediate and fast-acting (e.g., DeLong et al., 2014; Kuperberg & Jaeger, 2016). In Experiment 2, comprehenders were not observed to anticipate the two likely arguments of the verb when they encountered the verb plus voice information. In Experiment 3, the prediction of the unfolding sentential word order, expected to be demonstrated as a hypothesized cline of sentential word order patterns upon obtaining verb and voice morphology, was observed at a later region than expected. The hypothesized belief-update, after accruing information from the verb and the first noun phrase, was also demonstrated at a later sentence region. In Experiment 4, comprehenders demonstrated a diminishing ability to generate fine-grained predictions of the hypothesized gradience of sentential word order patterns when faced with time pressure. Experiment 5 replicated the findings from Experiment 3, and also showed how the contrastive effects are reduced in a reading task with memory load.

Overall, these findings indicate that successful predictive processing does not always occur and is sometimes not completed. Prediction outcomes can also be described in a continuum: the effects of predicting or anticipating the unfolding linguistic structure could be strong, weak, or non-existent, highlighting the gradient nature of predictive processes. The claim that successful prediction is limited contrasts with most of the extensive work on prediction (e.g., Altmann & Mirković, 2009; DeLong et al., 2014; Kuperberg & Jaeger, 2016), and aligns with claims that raise doubts about the ubiquity of successful predictions in real-time processing (e.g., Huettig, 2015; Huettig & Guerra, 2019; Luke & Christianson, 2016). Factors such as the amount of information available in the signal, as well as the quality or information types present in the linguistic context, may relate to the success of prediction outcomes. Both of these factors need to be examined in future research.

*Sentence parsing can be described as a dynamic process.* The evidence for the role of prediction in sentence processing has led several scholars to consider it as the core and unifying principle of the human mind (e.g., Bar, 2009; Clark, 2013; Friston, 2010; Lupyan & Clark, 2015). This dominance has resulted in less attention paid to other comprehension processes, such as integration. The degree of presence or absence of these effects were observed in the different experiments of this dissertation. In Experiment 2, an effect of anticipating the two likely arguments of the verb and voice was not demonstrated, nor was there a lack of differentiation across the four
voices as the first two sentential arguments unfolded. However, in Experiments 3 and 5, two different effects were observed at different time points, which captured how the unfolding linear word order of the sentence is processed in real time, i.e., a weak prediction effect and a strong integration effect of sentential word order. In Experiment 4, the addition of time pressure showed a reduction in the granularity of prediction, as well as the limits of the system to perform other relevant operations such as integration and belief-updating. These findings highlight how activations of the targeted linguistic representation at different time points vary as a function of the task.

This dissertation offered a dynamic account as a mechanism of sentence processing that explains the variability of successful prediction as observed from these experimental results. A dynamic account describes a cognitive-linguistic system that is gradient across multiple levels of representation and processes. The trajectories of changing activations and developing representations across levels are well-described by a dynamic account. I propose that both the cognitive processes of prediction and integration play significant roles in the changing activations of representations over the course of processing, where tipping points can describe the strength of a prediction to the developing representation. In the context of the dissertation results, the dynamic account is useful in explaining the range of predictive effects observed in processing, and how they relate to the amount of activation of the unfolding structure.
Appendix A
Guidelines for the Assessment of Correctness of Responses in Experiment 1

1. Overall, each response was analyzed for whether the correct nominal marker was used for every argument mentioned, relative to the verb.

<table>
<thead>
<tr>
<th>Voice Affix (Perfective)</th>
<th>Agent</th>
<th>Patient</th>
<th>Benefactive</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent [nag-]</td>
<td>ang</td>
<td>ng</td>
<td>para_sa</td>
<td>gamit_ang</td>
</tr>
<tr>
<td>Patient [-Ø-]</td>
<td>ng</td>
<td>ang</td>
<td>para_sa</td>
<td>gamit_ang</td>
</tr>
<tr>
<td>Benefactive [ipag-]</td>
<td>ng</td>
<td>ng</td>
<td>ang</td>
<td>gamit_ang</td>
</tr>
<tr>
<td>Instrument [ipang-]</td>
<td>ng</td>
<td>ng</td>
<td>para_sa</td>
<td>ang</td>
</tr>
</tbody>
</table>

Hence, the following are considered unacceptable (with the label codes in parentheses):

a. Marking an incorrect argument as the pivot, given the voice morphology (wrongpivot)
   (e.g., assigning a benefactive argument as the pivot when the voice morphology is in instrument voice)
   3PL
   *Ip<in>ang-hati yata kagabi… ng nanay ang cake para
   <PRF>IF-slice maybe last night NPVT mother PVVT cake for
   i-tabi sa anak niya na lalaki.
   PV.INF-set aside OBL child 3SG.NPVT REL boy
   ‘*Maybe the mother sliced with the cake last night (and) set (it) aside for her male child.’

   (P01, Item 7)

b. Marking two arguments as pivots within the matrix clause (doublepivot)

   *Ip<in>am-patay marahil kanina… ng apoy ang isa=ng baso
   <PRF>IF-put out perhaps a while ago NPVT fire PVVT one=LNK glass
   ng tubig ang bata ayon_sa utos ng kaniya=ng nanay.
   NPVT water PVVT child according to command PVVT 3SG,GEN=LNK mother
   ‘*Perhaps the child put out a glass of water the fire according to the mother’s command.’

   (P24, Item 50)

Note: The example below is acceptable (as discussed in Footnote 1) and is not considered doublepivot.
The absence of a pivot argument in the clause (no pivot argument) (**nopivot**)

*Ip<in>ag-putol noong isang buwan… ng lolo ang lola
<PRF>BF-cut the other month NPVT grandfather PVT grandmother
ng puno gamit_ang palakol.
NPVT tree INS axe
‘The other month, the grandfather cut the tree for the grandmother with an axe.’

Lack of an obligatory agent/patient argument (**noagent/nopatient**)

*Ip<in>am-basag noong makalawa… ng babae ng yelo
<PRF>IF-break two days ago NPVT womanNPVT ice
gamit_ang icepick para_sa bata.
INS icepick BEN child
‘*Two days ago, the woman broke with the ice the icepick for the child.’
(P22, Item 40)

Incorrect case markers for oblique arguments (**wrongmarkerins, wrongmarkerloc,** …)

*Ø-B<in>asag noong makalawa… ang yelo ng icepick ng nanay
PV-<PRF>break two days ago PVT ice NPVT icepick NPVT mother
para_sa ilagay sa inumin ng anak.
BEN INF-put LOC drink GEN child
‘*Two days ago, the mother broke the ice with the icepick for to be put in the child’s drink.’
(P09, Item 46)
Appendix B  
Guidelines for Syntactic Coding in Experiment 1  
(with examples from sample participant responses)

1. Each response was parsed in terms of the thematic roles of the arguments.

   Ip<in>ag-baliktad kaninang hapon… ng itlog ang lalaki  
   <PRF>BV-flip this afternoon NPVT egg PVT man  
   gamit_ang spatula para_sa bata.  
   INS spatula BEN child  
   ‘This afternoon, the man flipped the egg with the spatula for the child.’  
   Coding: PAT-AG-INS-BEN  
   (P11, Item 28)

   Ip<in>ag-putol noong isang buwan… ng lolo ang lola  
   <PRF>BF-cut the other month NPVT grandfather PVT grandmother  
   ng puno gamit_ang palakol.  
   NPVT tree INS axe  
   ‘The other month, the grandfather cut the tree for the grandmother with an axe.’  
   Coding: AG-BEN-PAT-INS  
   (P12, Item 29)

2. A complex noun phrase (italicized in this example) that referred to a single picture (in the visual scene) was coded as a single argument.

   Ip<in>an-tanggal kahapon… ng mantsa sa damit ng lalaki  
   <PRF>IV-remove yesterday NPVT stains LOC clothing NPVT man  
   ang sabon na Ø-g<in>amit ng babae.  
   PVT detergent REL PV<-PRF>use NPVT woman  
   ‘Yesterday, the man took off the stains on the clothing with the detergent the woman used.’  
   Coding: PAT-GEN-INS-[V-AG]  
   (P10, Item 50)

3. In cases where an argument can ambiguously be parsed as either an agent or genitive:

   a. An argument (italicized) was parsed as an agent if there is no other argument that can fulfill the obligatory agent role.

   Ip<in>an-tulak marahil… ng bangka ng lalaki ang kahoy  
   <PRF>IV-push perhaps NPVT boat NPVT man PVT wood  
   na i-b<in>igay ng lolo sa kanya.  
   REL PV<-PRF>give NPVT grandfather OBL 3SG.NPVT  
   ‘The man pushed the boat with the wood the grandfather gave to him.’
b. An argument was parsed as a genitive if it is followed by another argument that fulfills an agent role.

Nag-hubad yata kanina... ng sapatos ng bata ang kanya=ng ina. ‘Perhaps the mother removed earlier the child’s shoes.’

Coding: PAT-GEN-AG

(P10, Item 5)

In some cases, the agent need not necessarily follow the genitive.

Ip<a-mutol noong isang buwan... ng matanda=ng lalaki
<iv-cut last month NPVT old=lnk man
ang palakol sa puno ng isa=ng babae.
PVT axe OBL tree NPVT one=lnk woman
‘Last month, the old man cut with the axe the tree of a woman.’

Coding: AG-INS-PAT-GEN

(P01, Item 15)

4. Responses that were marked as OTHER (excluded from the data)

a. Those that involve other syntactic operations such as relativization or complementation

Ø-T<unaw yata kagabi... ang gamot na i-b<igay
PV-<melt perhaps last night PVT medicine REL PV-<give
ng doktor para sa pasyente
NPVT doctor BEN patient
‘The medicine given by the doctor for the patient was dissolved yesterday.’

Coding: PAT-[V-AG]-BEN

(P07, Item 49)

Ip<ag-pasan kani-kanina lang... ng pinya na naka-lagay sa tiklis
<BV-carry just a while ago NPVT pineapple REL STATIVE-placed OBL basket
ang lalaki para dalhin sa kanyang lola.
PVT man for bring-PV.INF OBL 3SG.GEN=LNK grandmother
‘Just a while ago, (someone) carried the pineapples placed in a basket for the man to be brought to his grandmother.’

Coding: PAT-[V-LOC]-BEN-[V-GOAL]

(P30, Item 15)
b. Those that only contain single NPs

\[ \text{Ø-K<in>} \text{atay kaninang tanghali... ang baboy} \]
\[ \text{PV-<PRF>} \text{slaughter at noon PVT pig} \]

‘(Someone) slaughtered the pig at noon.’
Coding: PAT

(P17, Item 15)

c. Those where the first or second arguments involved small clauses [SC]

\[ \text{Ip<in>} \text{ag-pasan kanina lang... ng basket na puno ng pinya} \]
\[ \text{<PRF>} \text{BV-carry just a while ago NPVT basket REL full NPVT pineapple} \]
\[ \text{ang lola na i-p<in>} \text{a-dala niya sa kanya=ng anak.} \]
\[ \text{PVT grandmother REL PV-CAUS<PRF>-send 3SG.NPVT OBL 3SG.GEN=LNK child} \]

‘Just a while ago, (someone) carried the basket full of pineapples for the grandmother who sent it to her child.’
Coding: PAT-[SC]-BEN-[V-AG-GOAL]

(P07, Item 35)

d. Those that involved some other non-relevant adjunct/prepositional phrase at the first or second argument position

\[ \text{Nag-basa noong isang araw... ang bata sa tulong ng nanay} \]
\[ \text{AV.PRF-read the other day PVT child OBL help GEN mother} \]
\[ \text{na g<um>} \text{a-gamit na ng salamin sa mata} \]
\[ \text{REL <AV>IIMP-use already NPVT eyeglasses LOC eyes} \]

‘The other day, the child read (something) with the help of the mother who was already using eyeglasses.’
Coding: AG-[PP]-[V-PAT-LOC]

(P01, Item 44)
Appendix C  
Stimuli Items in Experiment 1

This appendix contains the original list of stimuli items used in the study on Tagalog word order preferences, prior to exclusion or analysis. (Exclusions of verbs judged less grammatical and less natural were made prior to final analyses.) For every visual stimulus, the verb containing the voice morphology (in the agent, patient, benefactive, and instrumental voice, respectively) is presented, together with the adverbial phrase. Note that the order of these items do not reflect the order in which these were assigned or distributed across the experimental lists. All images in the visual scenes were taken from free clipart images online (Open Clip Art, 2017).

(1) Nagbaliktad / Binaliktad / Ipinagbaliktad / Ipinambaliktad kaninang hapon…

(2) Nagbanat / Binanat / Ipinagbanat / Ipinambanat yata kagabi…

(3) Nambasa / Binasa / Ipinagbasa / Ipinambasa yata kahapon…

(4) Nagbati / Binati / Ipinagbati / Ipinambati yata kanina…

(5) Nagbitbit / Binitbit / Ipinagbitbit / Ipinambitbit kahapon ng gabi…

(6) Nagpasan / Pinasan / Ipinagpasan / Ipinampasan kani-kanina lang…
(7) Nagpisa / Pinisa / Ipinagpisa / Ipinampisa kaninang tanghali…

(8) Nagsaway / Sinaway / Ipinagsaway / Ipinansaway kaninang umaga…

(9) Nagtulak / Tinulak / Ipinagtulak / Ipinantulak marahil kagabi…

(10) Nagyaya / Niyaya / Ipinagyaya / Ipinanyaya marahil kahapon…

(11) Nag-ahit / Inahit / Ipinag-ahit / Ipinang-ahit marahil kanina…

(12) Nag-alis / Inalis / Ipinag-alis / Ipinang-alis nang araw na iyon…

(13) Nag-ani / Inani / Ipinag-ani / Ipinang-ani noong isang araw…

(14) Nag-araro / Inararo / Ipinag-araro / Ipinang-araro noong isang buwan…

(15) Nag-ayos / Inayos / Ipinag-ayos / Ipinang-ayos noong isang gabi…

(16) Nagbalot / Binalot / Ipinagbalot / Ipinambalot noong isang linggo…
(17) Nagbasa / Binasa / Ipinagbasa / Ipinambasa noong isang taon…

(18) Nagbasag / Binasag / Ipinagbasag / Ipinambasag noong makalawa…

(19) Nagbilang / Binilang / Ipinagbilang / Ipinambilang siguro kagabi…

(20) Nagdurog / Dinurog / Ipinagdurog / Ipinandurog siguro kahapon…

(21) Naggupit / Ginupit / Ipinaggupit / Ipinanggupit siguro kanina…

(22) Naghabilin / Hinabilin / Ipinaghbilin / Ipinanghabilin kaninang hapon…

(23) Naghati / Hinati / Ipinaghati / Ipinanghati yata kagabi…

(24) Naghiwa / Hiniwa / Ipinaghiwa / Ipinanghiwa yata kahapon…

(25) Naghubad / Hinubad / Ipinaghubad / Ipinanghubad yata kanina…

(26) Nag-apon / Inipon / Ipinag-apon / Ipinang-apon kahapon ng gabi…
(27) Nagkahig / Kinahtig / Ipinagkahig / Ipinangkahig kani-kanina lang…

(28) Nagkatay / Kinatay / Ipinagkatay / Ipinangkatay kaninang tanghali…

(29) Nagkwenta / Kinwenta / Ipinagkwenta / Ipinangkwenta kaninang umaga…

(30) Naglaro / Nilaro / Ipinaglaro / Ipinanlaro marahil kagabi…

(31) Naglinis / Nilinis / Ipinaglinis / Ipinanlinis marahil kahapon…

(32) Nagpatay / Pinatay / Ipinapatay / Ipinampatay marahil kanina…

(33) Nagpiga / Piniga / Ipinapiga / Ipinampiga nang araw na iyon…

(34) Nagplantsa / Pinlantsa / Ipinagplantsa / Ipinamplantsa noong isang araw…

(35) Nagputol / Pinutol / Ipinagputol / Ipinamputol noong isang buwan…

(36) Nagsaing / Sinaing / Ipinagsaing / Ipinansaing noong isang gabi…
(37) Nagsayaw / Sinayaw / Ipinagsayaw / Ipinansayaw noong isang linggo…

(38) Nagsira / Sinira / Ipinagsira / Ipinansira noong isang taon…

(39) Nagsukat / Sinukat / Ipinagsukat / Ipinansukat noong makalawa…

(40) Nagsuklay / Sinuklay / Ipinagsuklay / Ipinansuklay siguro kagabi…

(41) Nagsunog / Sinunog / Ipinagsunog / Ipinansunog siguro kahapon…

(42) Nagtadtad / Tinadtad / Ipinagtadtad / Ipinantadtad siguro kanina…

(43) Nagtanggal / Tinanggal / Ipinagtanggal / Ipinantanggal kaninang hapon…

(44) Nagtunaw / Tinunaw / Ipinagtunaw / Ipinantunaw yata kagabi…

(45) Nagwalis / Winalis / Ipinagwalis / Ipinangwalis yata kahapon…

(46) Naghabi / Hinabi / Ipinaghabi / Ipinanghabi yata kanina…
(47) Naghanap / Hinanap / Ipinaghanap / Ipinanghanap kahapon ng gabi…

(48) Nagluto / Niluto / Ipinagluto / Ipinanluto kani-kanina lang…

(49) Nagsanay / Sinanay / Ipinagsanay / Ipinansanay kaninang tanghali…

(50) Nagsibak / Sinibak / Ipinagsibak / Ipinansibak kaninang umaga…

(51) Nagtahi / Tinahi / Ipinagtahi / Ipinantahi marahil kagabi…

(52) Nagpinta / Pininta / Ipinagpinta / Ipinampinta marahil kanina…

Reference
Appendix D
Stimuli Items and Comprehension Questions in Experiment 2

The visual stimuli for both critical and filler items can be accessed in my OSF account:
(https://osf.io/n5tjy/)

Critical Items
In the English translations of the Tagalog sentences, the adverb of uncertainty (present in selected items) is placed before the verb and the time adverb is placed at the end of the sentence. In the Tagalog sentences, both adverbs occur after the verb and before the first noun phrase.

1. banat (stretch)
The man perhaps stretched the spring for the child with the pliers last night.
   AV: Nagbanat yata kagabi ang lalaki ng spring para sa bata gamit ang plais.
   Nagbanat ba ang tatay ng spring? (Did the father stretch the spring?)
   PV: Binanat yata kagabi ng lalaki ang spring para sa bata gamit ang plais.
   Binanat ba ng tatay ang spring? (Did the father stretch the spring?)
   BV: Ipinagbanat yata kagabi ng lalaki ang bata ng spring gamit ang plais?
   Ipinagbanat ba ng tatay ang bata? (Did the father stretch for the child?)
   IV: Ipinangbanat yata kagabi ng lalaki ang plais ng spring para sa bata.
   Ipinambanat ba ng tatay ang plais? (Did the father stretch with the pliers?)

2. bitbit (carry (with the hands))
The man carried the vegetables for the woman with the cart yesterday evening.
   AV: Nagbitbit kahapon ng gabi ang lalaki ng gulay para sa babae gamit ang kariton.
   Nagbitbit ba ang lalaki para sa babae? (Did the man carry for the woman?)
   PV: Binitbit kahapon ng gabi ng lalaki ang gulay para sa babae gamit ang kariton.
   Binitbit ba ng lalaki ang gulay? (Did the man carry the vegetables?)
   BV: Ipinagbitbit kahapon ng gabi ng lalaki ang bata ng gulay gamit ang kariton.
   Ipinagbitbit ba ang babae ng mangoes? (Did the man carry the vegetables?)
   IV: Ipinangbitbit kahapon ng gabi ng lalaki ang kariton nga gulay para sa babae.
   Ipinambitbit ba ng lalaki ang kariton? (Did the man carry with the cart?)

3. pasan (carry (over the shoulders))
The son carried the pineapple for the grandmother with the container just a while ago.
   AV: Nagpasan kani-kanina lang anak ng pinya para sa lola gamit ang basket.
   Nagpasan ba ang anak gamit ang bilao? (Did the son carry with the flat basket?)
   PV: Pinasan kani-kanina lang ng anak ang pinya para sa lola gamit ang basket.
   Pinasan ba ng anak ang mangga? (Did the son carry the mangoes?)
   BV: Ipinagpasan kani-kanina lang ng anak ang lola ng pinya gamit ang basket.
   Ipinagpasan ba ng anak ang tatay? (Did the son carry for the father?)
   IV: Ipinangpasan kani-kanina lang ng anak ang basket nga pinya para sa lola.
   Ipinampasan ba ng anak ang bilao? (Did the son carry with the flat basket?)

4. tulak (push)
The young man pushed the boat for the elderly with the wood this afternoon.
   AV: Nagtulak kaninang tanghali ang binata ng bangka para sa matanda gamit ang kahoy.
   Nagtulak ba ang babae gamit ang kahoy? (Did the woman push with the wood?)
PV: Tinulak kaninang tanghali ng binata ang bangka para sa matanda gamit ang kahoy. 
   Tinulak ba ng babae ang bangka? (Did the woman push the boat?)
BV: Ipinagtulak kaninang tanghali ng binata ang matanda ng bangka gamit ang kahoy.
   Ipinagtulak ba ng babae ang matanda? (Did the woman push for the elderly?)
IV: Ipinantulak kaninang tanghali ng binata ang kahoy ng bangka para sa matanda.
   Ipinantulak ba ng babae ang kahoy? (Did the woman push with the wood?)

5. ahit (shave)
The barber perhaps shaved the mustache for the man with the razor earlier.
AV: Nagahit marahil kanina ang barbero ng bigote para sa lolo gamit ang labaha.
   Nag-ahit ba ang barbero ng bigote? (Did the barber shave the mustache?)
PV: Inahit marahil kanina ng barbero ang bigote para sa lolo gamit ang labaha.
   Inahit ba ng barbero ang bigote? (Did the barber shave the mustache?)
BV: Ipinahit marahil kanina ng barbero ang lolo ng bigote gamit ang labaha.
   Ipinahit ba ng barbero ang lolo? (Did the barber shave for the grandfather?)
IV: Ipinangahit marahil kanina ng barbero ang labaha ng bigote para sa lolo.
   Ipinangahit ba ng barbero ang labaha? (Did the barber shave with the razor?)

6. ani (harvest)
The young man harvested the mango for the elderly (man) with the pole/stick the other day.
AV: Nag-ani noong isang araw ang binata ng mangga para sa matanda gamit ang sungkit.
   Nag-ani ba ang binata ng bayabas? (Did the young man harvest the guava?)
PV: Inani noong isang araw ng binata ang mangga para sa matanda gamit ang sungkit.
   Inani ba ng binata ang bayabas? (Did the young man harvest the guava?)
BV: Ipinagani noong isang araw ng binata ng mangga gamit ang sungkit.
   Ipinag-ani ba ng bayabas ang matanda? (Was the guava harvested for the elderly?)
IV: Ipinangani noong isang araw ang sungkit ng mangga para sa matanda.
   Ipinang-ani ba ng bayabas ang sungkit? (Was the guava harvested with the pole?)

7. araro (plow)
The farmer plowed the farm for the businessman with the hoe last month.
AV: Nag-araro noong isang buwan ang magsasaka ng bukid para sa negosyante gamit ang asarol.
   Nag-araro ba ang karpintero ng bukid? (Did the carpenter plow the farm?)
PV: Inararo noong isang buwan ng magsasaka ang bukid para sa negosyante gamit ang asarol.
   Inararo ba ng karpintero ang bukid? (Did the carpenter plow the farm?)
BV: Ipinagararo noong isang buwan ng magsasaka ang negosyante ng bukid gamit ang asarol.
   Ipinag-araro ba ng karpintero ang negosyante? (Did the carpenter plow for the businessman?)
IV: Ipinangagararo noong isang buwan ng magsasaka ang asarol ng bukid para sa negosyante.
   Ipinang-araro ba ng karpintero ang asarol? (Did the carpenter plow with the hoe?)

8. balot (cover, wrap)
The mother wrapped the book for the child with the plastic cover last week.
AV: Nagbalot noong isang linggo ang nanay ng aklat para sa bata gamit ang plastik.
Nagbalot ba ang nanay gamit ang plastik? (Did the mother wrap with the plastic cover?)
PV: Binalot noong isang linggo ng nanay ang aklat para sa bata gamit ang plastik.  
Binalot ba ng nanay ang libro? (Did the mother wrap the book?)
BV: Ipinagbalot noong isang linggo ng nanay ang aklat para sa bata gamit ang plastik.  
Ipinagbalot ba ng nanay ang bata? (Did the mother wrap for the child?)
IV: Ipinangbalot noong isang linggo ng nanay ang plastik ng aklat para sa bata.  
Ipinambalot ba ng nanay ang plastik? (Did the mother wrap with the plastic cover?)

9. basag (break)
The woman broke the ice for the child with the ice pick the day before yesterday.
AV: Nagbasag noong makalawa ang babae ng yelo para sa anak gamit ang ice pick.  
Nagburo ba ang nanay ng yelo? (Did the mother form the ice?)
PV: Binasag noong makalawa ng babae ang yelo para sa anak gamit ang ice pick.  
Bino ng nanay ang yelo? (Did the mother form the ice?)
BV: Ipinagbasag noong makalawa ng babae ang anak ng yelo gamit ang ice pick.  
Ipinagburo ba ng nanay ang yelo ang anak? (Did the mother form for the child?)
IV: Ipinangbasag noong makalawa ng babae ang ice pick ng yelo para sa anak.  
Ipinambuog ba ng nanay ang yelo? (Did the mother form for the child?)

10. bilang (count)
The woman perhaps counted the strawberry for the man with the basket last night.
AV: Nagbilang siguro kagabi ang babae ng strawberry para sa lalaki gamit ang basket.  
Nagbilang ba ang nanay ng strawberry? (Did the woman count the strawberry?)
PV: Binilang siguro kagabi ng babae ang strawberry para sa lalaki gamit ang basket.  
Binilang ba ng nanay ang strawberry? (Did the woman count the strawberry?)
BV: Ipinagbilang siguro kagabi ng babae ang lalaki ng strawberry gamit ang basket.  
Ipinagbilang ba ng lalaki ang lalaki? (Did the woman count for the man?)
IV: Ipinangbilang siguro kagabi ng babae ang basket ng strawberry para sa lalaki.  
Ipinambilang ba ng strawberry ang basket? (Was the strawberry counted with the basket?)

11. durog (pound)
The mother perhaps mashed the carrot for the child with the spoon yesterday.
AV: Nagdurog siguro kahapon ang nanay ng karot para sa anak gamit ang kutsara.  
Nagdurog ba ang nanay ng patatas? (Did the mother mash the potato?)
PV: Dinurog siguro kahapon ng nanay ang karot para sa anak gamit ang kutsara.  
Dinurog ba ng nanay ang patatas? (Did the mother mash the potato?)
BV: Ipinagdurog siguro kahapon ng nanay ang anak ng karot gamit ang kutsara.  
Ipinagdurog ba ng patatas ang anak? (Did the mother mash for the child?)
IV: Ipinangdurog siguro kahapon ng nanay ang kutsara ng karot para sa anak.  
Ipinandurog ba ng patatas ang kutsara? (Was the potato mashed with the spoon?)

12. gupit (cut)
The nurse perhaps cut the bandage for the patient with the scissors earlier.
AV: Naggupit siguro kanina ang nars ng benda para sa pasyente gamit ang gunting.  
Naggupit ba ang nars ng benda? (Did the nurse cut the bandage?)
PV: Ginupit siguro kanina ng nars ang benda para sa pasyente gamit ang gunting.  
Ginupit ba ng nars ang benda? (Did the nurse cut the bandage?)
15. hiliwa (slice)
The chef perhaps sliced the watermelon for the woman with the knife yesterday.

AV: Naghiwa yata kahapon ang kusinero ng pakwan para sa babae gamit ang kutsilyo.
   Naghiwa ba ang kusinero ng patola? (Did the chef slice the gourd?)

PV: Hiniwa yata kahapon ng kusinero ang pakwan para sa babae gamit ang kutsilyo.
   Hiniwa ba ng kusinero ang patola? (Did the chef slice the gourd?)

BV: Ipinaghiwa yata kahapon ng kusinero ang pakwan para sa babae gamit ang kutsilyo.
   Ipinaghiwa ba ng kusinero ang patola? (Was the gourd sliced with the knife?)

16. kahig (scratch (on the ground))
The chicken scratched the rice grains for the chick with the claw just a while ago.

AV: Nagkahig kani-kanina lang ang manok ng palay para sa sisiw gamit ang kuko.
   Nagkahig ba ang manok para sa sisiw? (Did the chicken scratch for the chick?)

PV: Kinahig kani-kanina lang ng manok ang palay para sa sisiw gamit ang kuko.
   Kinahig ba ng manok ang palay? (Did the chicken scratch the rice grains?)

BV: Ipinagkahig kani-kanina lang ng manok ang sisiw ng palay gamit ang kuko.
   Ipinagkahig ba ng manok ang sisiw? (Did the chicken scratch for the chick?)

IV: Ipinangkahig kani-kanina lang ng manok ang kuko ng palay para sa sisiw.
   Ipinangkahig ba ng manok ang kuko? (Did the chicken scratch with the claw?)

17. katay (kill (an animal))
The vendor butchered the pork for the customer with the knife this afternoon.
AV: Nagkatay kaninang tanghali ang tindera ng baboy sa mamimili gamit ang itak.
PV: Kinatay kaninang tanghali ng tindera ang baboy para sa mamimili gamit ang itak.
BV: Ipinagkatay kaninang tanghali ng tindera ang mamimili ng baboy gamit ang itak.
IV: Ipinangkatay kaninang tanghali ng tindera ang itak na baboy para sa mamimili.

18. kwenta (compute)
The accountant computed the money for the grandmother with the calculator this morning.
AV: Nagkwenta kaninang umaga ang accountant ng pera para sa lola gamit ang calculator.
PV: Kinwenta kaninang umaga ng accountant ang pera para sa lola gamit ang calculator.
BV: Ipinagkwenta kaninang umaga ng accountant ang lola ng pera gamit ang calculator.
IV: Ipinangkwenta kaninang umaga ng accountant ang pera para sa lola.

19. linis (clean, wash)
The man perhaps cleaned the car for the police with the sponge yesterday.
AV: Naglinis marahil kahapon ang lalaki ng kotse para sa pulis gamit ang espongha.
PV: Nilinis marahil kahapon ng lalaki ang kotse para sa pulis gamit ang espongha.
BV: Ipinaglinis marahil kahapon ng lalaki ang pulis ng kotse gamit ang espongha.
IV: Ipinanglinis marahil kahapon ng lalaki ang espongha ng kotse para sa pulis.

20. patay (kill, put out)
The mother perhaps put off the fire for the child with the water earlier.
AV: Nagpatay marahil kanina ang nanay ng apoy para sa anak gamit ang tubig.
PV: Pinatay marahil kanina ng nanay ang apoy para sa anak gamit ang tubig.
BV: Ipinagpatay marahil kanina ng nanay ang apoy gamit ang tubig.

21. piga' (squeeze)
The mother squeezed the lime for the child with the strainer that day.
AV: Nappiga' nang araw na iyong ang nanay ng kalamansi para sa anak gamit ang salaan.
Nagpiga ba ang tatay ng kalamansi? (Did the father squeeze the lime?)
PV: Piniga' nang araw na iyon ng nanay ang kalamansi para sa anak gamit ang salaan.
Piniga ba ng tatay ang kalamansi? (Did the father squeeze the lime?)
BV: Ipinagpiga' nang araw na iyon ng nanay ang anak ng kalamansi gamit ang salaan.
Ipinagpiga ba ng tatay ang anak? (Did the father squeeze the child?)
IV: Ipinangpiga' nang araw na iyon ng nanay ang salaan ng kalamansi para sa anak.
Ipinampiga ba ng tatay ang strainer? (Did the father squeeze with the strainer?)

22. plantsa (iron (the clothes))
The woman ironed the shirt for the child with the iron board the other day.
AV: Nagplantsa noong isang araw ang babae ng damit para sa bata gamit ang kabayo.
Nagsampay ba ang babae ng damit? (Did the woman hang the clothes?)
PV: Pinlantsa noong isang araw ng babae ang damit para sa bata gamit ang kabayo.
Sinampay ba ng babae ang damit? (Did the woman hang the clothes?)
BV: Ipinagplantsa noong isang araw ng babae ang bata ng damit gamit ang kabayo.
Ipinagsampay ba ng babae ang bata? (Did the woman hang for the child?)
IV: Ipinangplantsa noong isang araw ng babae ang kaldero ng damit para sa bata.
Ipinamgsampay ba ng babae ang kabayo? (Did the woman hang with the iron board?)

23. putol (break, cut)
The grandfather cut the tree for the grandmother with the axe last month.
AV: Nagputol noong isang buwan ang lolo ng puno para sa lola gamit ang palakol.
Nilagari ba ng lolo para sa lola? (Did the grandfather saw for the grandmother?)
PV: Pinutol noong isang buwan ng lolo ang puno para sa lola gamit ang palakol.
Nilagari ba ng lolo ang puno? (Did the grandfather saw the tree?)
BV: Ipinagputol noong isang buwan ng lolo ang lola ng puno gamit ang palakol.
Ipinaglagari ba ng lolo ang lola? (Did the grandfather saw for the grandmother?)
IV: Ipinangputol noong isang buwan ng lolo ang palakol ng puno para sa lola.
Ipinanlagari ba ng lolo ang palakol? (Did the grandfather saw with the axe?)

24. sa'ing (cook rice)
The father cooked rice for the child with the cooking pot last night.
AV: Nagsa'ing noong isang gabi ang tatay ng kanin para sa bata gamit ang kaldero.
Nagsaing ba ang tatay ng kanin? (Did the father cook rice?)
PV: Sina'ing noong isang gabi ang tatay ng kanin para sa bata gamit ang kaldero.
Sinaing ba ang tatay ng kanin? (Did the father cook rice?)
BV: Ipinagsa'ing noong isang gabi ang bata ng kanin gamit ang kaldero.
Ipinagsaing ba ang tatay ang bata? (Did the father cook for the child?)
IV: Ipinangsa'ing noong isang gabi ang kaldero ng kanin para sa bata.
Ipinanlsaing ba ng tatay ang kaldero? (Did the father cook with the cooking pot?)

25. sayaw (dance)
The child danced (to) the music for the grandfather with the shoes last week.
AV: Nagsayaw noong isang linggo ang bata ng tugtugin para sa lolo gamit ang sapatos.
Nagsayaw ba ang lolo ng tugtugin? (Did the grandfather dance (to) the music?)
PV: Sinayaw noong isang linggo ang bata ang tugtugin para sa lolo gamit ang sapatos.
Sinayaw ba ang lolo ang tugtugin? (Did the grandfather dance (to) the music?)
BV: Ipinagsayaw noong isang linggo ng bata ang lolo ng tugtugin gamit ang sapatos.
Ipinagsayaw ba ng lolo ang bata? (Did the grandfather dance for the child?)
IV: Ipinansayaw noong isang linggo ng bata ang sapatos ng tugtugin para sa lolo. (Did the grandfather dance with the shoes?)

26. suklay (comb)
The mother perhaps combed the hair for the sibling with the hairbrush last night.
AV: Nagsuklay siguro kagabi ang nanay ng buhok para sa bata gamit ang hairbrush. (Did the mother cut hair?)
PV: Sinuklay siguro kagabi ang nanay ng buhok para sa bata gamit ang hairbrush. (Did the mother cut hair?)
BV: Ipinagsuklay siguro kagabi ng nanay ang buhok gamit ang hairbrush. (Did the mother cut for the child?)
IV: Ipinansuklay siguro kagabi ng nanay ang hairbrush ng buhok para sa bata. (Did the mother cut with the comb?)

27. tadtad (chop)
The chef perhaps chopped the meat for the woman with the knife earlier.
AV: Nagtadtad siguro kanina ang kusinero ng karne para sa babae gamit ang kutsilyo. (Did the chef chop with the knife?)
PV: Tinadtad siguro kanina ng kusinero ang karne para sa babae gamit ang kutsilyo. (Did the chef chop the meat?)
BV: Ipinagtadtad siguro kanina ng kusinero ang babae ng karne gamit ang kutsilyo. (Did the chef chop for the woman?)
IV: Ipinangadtad siguro kanina ng kusinero ang kutsilyo ng karne para sa babae. (Did the chef chop with the comb?)

28. tanggal (take (something) out)
The woman washed the stains for the man with the detergent this afternoon.
AV: Nagtanggal kaninang hapon ang babae ng mantsa para sa lalaki gamit ang sabon. (Did the woman wash the stain?)
PV: Tinanggal kaninang hapon ng babae ang mantsa para sa lalaki gamit ang sabon. (Did the woman wash the stain?)
BV: Ipinagtanggal kaninang hapon ng babae ang mantsa para sa lalaki gamit ang sabon. (Was the stain washed for the man?)
IV: Ipinangtanggal kaninang hapon ng babae ang sabon ng mantsa para sa lalaki. (Were the stains washed with the soap?)

29. tunaw (melt)
The doctor perhaps dissolved the medicine for the patient with the water yesterday.
AV: Nagtunaw yata kagabi ang doktor ng gamot para sa pasyente gamit ang tubig. (Did the doctor extract medicine?)
PV: Tinunaw yata kagabi ng doktor ang gamot para sa pasyente gamit ang tubig. (Did the doctor extract for the patient?)
BV: Ipinagtunaw yata kagabi ng doktor ang pasyente ng gamot gamit ang tubig. (Did the doctor extract with the water?)
IV: Ipinangtunaw yata kagabi ng doktor ang pasyente ng gamot gamit ang tambo. (Did the doctor extract the medicine with the broom stick?)

30. walis (sweep)
The maid perhaps swept the leaves for the boss with the broom stick yesterday.
AV: Nagwalis yata kahapon ang nanay ng dahon para sa tatay gamit ang tambo.
31. habi' (weave)

The woman perhaps weaved the quilt for the daughter with the string earlier.

AV: Naghabi' yata kanina ang nanay ng kumot para sa anak gamit ang sinulid.
Nagbuo ba ang nanay ng sinulid? (Did the woman form the thread (ball)?)

PV: Nilahod kaniyana yata ng nanay ang kumot para sa anak gamit ang sinulid.
Binuo ba ng nanay ang sinulid? (Did the woman form the thread (ball)?)

BV: Ipinaghabi' yata kanina ng nanay ang anak ng kumot gamit ang sinulid.
Ipinagbuo ba ng nanay ang anak? (Was the thread (ball) formed for the child?)

IV: Ipinambuo kanina lang ang nanay ang kawali ng kumot para sa anak.
Ipinambuo ba ng sinulid ang kawali? (Was the thread (ball) formed with the wood?)

32. hanap (search)

The father searched the ring for the wife with the flashlight yesterday evening.

AV: Naghanap kahapon ng gabi ang tatay ng sising para sa babae gamit ang flash light.
Hinanap ba ang tatay ng sising? (Did the man search for the ring?)

PV: Hinahap kahapon ng gabi ng tatay ang sising para sa babae gamit ang flash light.
Hinanap ba ng talaga ang sising? (Did the man search for the ring?)

BV: Ipinaghanap kahapon ng gabi ang babae ng sising gamit ang flash light.
Ipinaghanap ba ng sising ang asawang babae? (Was the ring sought for the wife?)

IV: Ipinanghanap kahapon ng gabi ang flash light ng sising para sa babae.
Ipinanghanap ba ng sising ang flash light? (Was the ring sought with the flashlight?)

33. luto' (cook)

The mother cooked the spaghetti for the child with the pan just a while ago.

AV: Nagluto' kaniyana ang nanay ng spaghetti para sa anak gamit ang kawali.
Nagluto ba ang nanay ng spaghetti? (Did the mother cook spaghetti?)

PV: Niluto' kaniyana lang ang nanay ang spaghetti para sa anak gamit ang kawali.
Niluto ba ang nanay ang spaghetti? (Did the mother cook spaghetti?)

BV: Ipinagluto' kaniyana ang nanay ang anak ng spaghetti gamit ang kawali.
Ipinagluto ba ng nanay ang anak? (Did the mother cook for the child?)

IV: Ipinangluto' kaniyana lang ang nanay ang kawali ng spaghetti para sa anak.
Ipinangluto ba ng nanay ang kawali? (Did the mother cook with the pan?)

34. sibak (chop (wood))

The woodcutter chopped the wood for the mother with the axe this morning.

AV: Nagsibak kaninang umaga ang talaga ng kahoy para sa babae gamit ang palakol.
Nagsibak ba ang talaga ang kawali? (Did the man chop with the saw?)

PV: Sinibak kaninang umaga ng talaga ang kawali gamit ang palakol.
Sinibak ba ng talaga ang mga sanga? (Did the man chop the branches?)

BV: Ipinagsibak kaninang umaga ng talaga ang babae ng kawali gamit ang palakol.
Ipinagsibak ba ng mga sanga ang babae? (Were the branches chopped for the woman?)

IV: Ipinangsibak kaninang umaga ng lalaki ang palakol ng kahoy para sa babae.

Ipinansibak ba ng sanga ang lagari? (Were the branches chopped with the saw?)

35. tahi' (sew)
The grandmother perhaps sewed the dress for the grandchild with the needle last night.

AV: Nagtahi' marahil kagabi ang Lola ng damit para sa apo gamit ang karayom.

PV: Tinahi' marahil kagabi ang Lola ng damit para sa apo gamit ang karayom.

BV: Ipinagtahi' marahil kagabi ang Lola ng damit para sa apo gamit ang karayom.

Ipinagtahi ba ng sastre ang damit? (Did the tailor sew the dress?)

36. pinta (paint)
The student perhaps painted the picture for the teacher with the paint brush earlier.

AV: Nagpinta marahil kagabi ang estudyante ng larawan para sa guro gamit ang brush.

PV: Pininta marahil kagabi ang estudyante ng larawan para sa guro gamit ang brush.

BV: Ipinagpinta marahil kagabi ang estudyante ng larawan para sa guro gamit ang brush.

Filler Items
In the English translations, the adverb of uncertainty (present in selected items) is placed before the verb and the time adverb is placed after one of the nominal elements. In the Tagalog sentences, both adverbs occur after the verb and before the first noun phrase.

Unergative Intransitives
The child perhaps cried earlier to the teacher in the classroom because of the exam.

2. Tumahimik noong isang araw ang pusa sa daan dahil sa aso mula sa kulungan.
The cat became silent on the street the other day because of the dog from the cage.

3. Umurong noong isang gabi ang babae mula sa kotse dahil sa pulis sa kanto.
The girl returned the other night from the car because of the policeman at the corner.

The old man made up last year for the competition against the young man for the trophy.

5. Bumalik siguro kagabi ang estudyante sa paaralan mula sa tindahan ayon sa janitor.
The student perhaps returned last night to school from the store according to the janitress.
Bumalik ba ang estudyante sa aklatan? (Did the student return to the library?)

6. Tumuloy siguro kanina ang babae sa silungan dahil sa ulan ayon sa bata.
   The woman perhaps headed earlier to the shed because of the rain according to the child.
   Tumuloy ba ang lalaki sa silungan? (Did the man head to the shed?)

7. Sumama yata kagabi ang anak sa tatay mula sa laru:an hanggang sa simbahan.
   The child perhaps joined the father last night from the playground to the church.
   Sumama ba ang anak sa tatay? (Did the child join the father?)

   The man perhaps gasped (for breath) earlier in the garden because of the dog according to the girl.
   Humingal ba ang lalaki sa bahay? (Did the man gasp (for breath) at home?)

   The man pretended the man as a clown in the circus for the child just a while ago.
   Nagkunwari ba ang lalaki bilang clown? (Did the man pretend as a clown?)

    The rabbit disappeared this morning in the box because of the magician according to the child.
    Naglaho ba ang kuneho ayon sa bata? (Did the rabbit disappear according to the child?)

    The teacher perhaps spoke up yesterday in the classroom regarding the books for the student.
    Nagsalita ba ang guro tungkol sa pagsusulit? (Did the teacher speak about the exam?)

    The mother doubted the man that day because of the package from the airport.
    Nagtaka ba ang nanay dahil sa bagahe? (Did the mother doubt because of the luggage?)

    The child threw (tantrums) in the store last month because of the toy according to the grandmother.
    Nagwala' ba ang bata dahil sa laruan? (Did the child throw (tantrums) because of the toy?)

    The man caused (a rumble) on the street last week because of the motorist despite the traffic.
    Naghuramentado ba ang lalaki sa kalsada? (Did the man cause (a rumble) on the street?)

    The actor became poor the day before yesterday because of debt from the casino according to the spouse.
    Naghirap ba ang aktor ayon sa asawa? (Did the actor become poor according to the spouse?)

    The students perhaps discussed yesterday about the money as help for the beggar.

163
Nag-usig ba ang mga estudyante ukol sa pera? (Did the students inquire about the money?)

17. Tumalino noong isang buwan ang bata dahil sa sudoku mula sa libro ayon sa guro.
   The child became intelligent last month because of sudoku from the books according to the teacher.
   Tumalino ba ang guro dahil sa sudoku? (Did the teacher become intelligent because of sudoku?)

   The woman perhaps became cranky with the man last night because of the call from the restaurant.
   Sumungit ba ang dalaga sa binata? (Did the woman become cranky with the man?)

   The rain perhaps weakened earlier because of the spell from the witch according to the hermit.
   Humina' ba ang bagyo dahil sa kulam? (Did the typhoon weaken because of the spell?)

   The woman became rich with the business just a while ago because of the donation from the composer.
   Yumaman ba ang dalaga sa donasyon? (Did the woman become rich because of the donation?)

   The student became industrious at school this morning because of the gift from the father.
   Sumipat ba ang estudyante sa paaralan? (Did the student leave school?)

22. Nabingi marahil kanina ang lolo dahil sa ingay mula sa karaoke ayon sa anak.
   The grandfather perhaps became deaf earlier because of the noise from the karaoke according to the child.
   Nabingi ba ang lolo sa ingay? (Did the grandfather become deaf with the noise?)

   The businessman perhaps lost (profit) from the store last night because of the fire according to the captain.
   Nalugi ba ang negosyante ayon sa kapitan? (Did the businessman lose profit according to the captain?)

Transitives with Experiencer Pivots

   The children watched the animals this afternoon in the television according to the mother.
   Nanood ba ang mga bata ayon sa nanay? (Did the children watch according to the mother?)

   The girl perhaps listened to the song on the radio yesterday according to the mother.
   Nakinig ba ang lalaki ng awit? (Did the man listen to the song?)

The mechanic came from the repair shop besides the store yesterday evening because of the customer.
Nanggaling ba ang mekaniko sa pabrika? (Did the mechanic come from the repair shop?)

27. Nanggigil kaninang tanghali ang lola sa sanggol dahil sa braso bukod pa sa damit.
The woman found the baby (cute) because of (her) arms other than (her) clothes this afternoon.
Nanggigil ba ang lola sa sanggol? (Did the grandmother find the baby cute?)

28. Lumigaya marahil kagabi ang guro sa plorera bilang regalo mula sa estudyante.
The teacher perhaps became happy of the vase last night (which served) as a gift from the student.
Lumigaya ba ang estudyante sa plorera? (Did the teacher become happy of the vase?)

29. Sumaya yata kagabi ang babae sa bulaklak bukod pa sa tsokolate buhat sa kasintahan.
The girl perhaps became happy for the flowers aside from the chocolate last night from the boyfriend.
Sumaya ba ang babae sa bulaklak? (Did the girl become happy with the flowers?)

The worker perhaps suffered earlier from because of the things (to be carried) in the mountain.
Nagbasa ba ang manggagawa sa bundok? (Did the worker read in the mountain?)

The woman got upset with the man just a while ago because of the call from the office.
Nagtampo ba ang babae dahil sa text? (Did the woman get upset because of the text?)

32. Naawa kaninang umaga ang lolo sa musikero dahil sa pananamit bukod pa sa pagkain.
The grandfather pitied the musician this morning because of (his) clothes besides (his) food.
Naawa ba ang lolo sa musikero? (Did the grandfather pity the musician?)

33. Nabuhay marahil kahapon ang halaman dahil sa insecticide mula sa sayantipiko ayon sa magsasaka.
The plant perhaps became alive yesterday because of the insecticide from the scientist according to the farmer.
Nabuhay ba ang halaman dahil sa insecticide? (Did the plant become alive because of the insecticide?)

34. Nalungkot nang araw na iyon ang estudyante dahil sa eksam mula sa punung-guro sa paaralan.
The student got saddened that day because of the exam from the principal in the school.
Nalungkot ba ang guro dahil sa eksam? (Did the teacher get saddened because of the exam?)

35. Nanginig noong isang buwan ang bata dahil sa multo mula sa kwarto ayon sa lolo.
The child shivered last month because of the ghost inside the room according to the grandfather.
Nanginig ba ang bata dahil sa hayop? (Did the child shiver because of the animal?)

36. Natuwang isang linggo ang bata sa sorbetes bukod pa sa aginaldo galing sa estranghero.
The grandchild became happy last week because of the ice cream besides the gift from the stranger.
Natuwa' ba ang bata sa aginaldo? (Did the child become happy with the gift?)

37. Nauhaw noong makalawa ang siklista dahil sa paligsahan sa bundok ayon sa nobya.
   The cyclist felt thirsty the day before yesterday because of the competition in the
   mountain according to the girlfriend.
   Nauhaw ba ang siklista ayon sa nobya? (Did the cyclist feel thirsty according to the
   girlfriend?)

38. Namanhid siguro kahapon ang mga paa ng lolo dahil sa gamot mula sa nars.
   The feet perhaps felt numb yesterday according to the grandfather because of the
   medicine from the nurse.
   Namanhid ba ang mga paa dahil sa gamot? (Did the feet feel numb because of the
   medicine?)

   The man felt numb this afternoon because of the things from the house according to the
   spouse.
   Nangawit ba ang lalaki dahil sa mga gamit? (Did the man feel numb because of the
   things?)

40. Nangilabot yata kahapon ang babae sa ulan bukod pa sa kidlat ayon sa estranghero.
   The woman perhaps felt scared with the rain yesterday besides the thunder according to
   the stranger.
   Nangilabot ba ang lalaki dahil sa kidlat? (Did the woman feel scared because of the
   thunder?)

41. Nanabik kahapon ng gabi ang bata sa kendi bukod pa sa laruan mula sa ninang.
   The child became excited with the candies besides the toy from the godmother yesterday
   evening.
   Nanalo ba ang bata ng kendi? (Did the child win the candy?)

42. Nanliit kaninang tanghali ang tuta sa silong dahil sa baka mula sa bukid.
   The puppy under the shed felt intimidated this afternoon with the cow from the farm.
   Nanliit ba ang tuta sa baka? (Did the puppy feel intimidated with the cow?)

43. Inalam noong isang taon ng bata ang daan mula sa babae ayon sa palatandaan.
   The child aimed to know the road based on the sign from the woman last year.
   Inalam ba ng babae ang daan? (Did the woman aim to know the road?)

44. Kinaya siguro kagabi ng pasyente ang hagdan sa ospital ayon sa aktor.
   The player perhaps managed (to climb) the stairs in the hospital last night according to
   the actor.
   Kinaya ba ng pasyente ang hagdan? (Did the patient manage the stairs?)

45. Pinalad siguro kanina ang manlalaro sa paligsahan sa telebisyon ayon sa aktor.
   The player perhaps became fortunate in the competition earlier in the television according
   to the actor.
   Pinalad ba ang manlalakbay sa paligsahan? (Was the traveler fortunate with the
   competition?)

46. Sinulit yata kagabi ng lalaki ang bakasyon sa dalampasigan alang-alang sa kasintahan.
   The man perhaps maximized the vacation in the beach last night for the girlfriend.
   Sinulit ba ang lalaki ang bakasyon? (Did the man maximize the vacation?)

47. Sinwerte yata kanina ang babae sa roleta dahil sa dasal mula sa pari.
The woman perhaps became lucky in the roulette earlier because of the prayer from the priest.
Sinwerte ba ang pari sa rolleta? (Did the priest become lucky with the roulette?)

Unaccusative Intransitives
The truck perhaps ran yesterday because of the mechanic from the factory according to the man.
Umandar ba ang trak dahil sa mekaniko? (Did the truck run because of the mechanic?)
49. Gumaling marahil kanina ang sugat dahil sa gamot mula sa doktor ayon sa pasyente.
The wound perhaps healed yesterday because of the medicine from the doctor according to the patient.
Lumala ba ang sugat ayon sa doktor? (Did the wound get worse according to the doctor?)
50. Umilaw noong isang araw ang bombilya dahil sa baterya mula sa elektrisista ayon sa kustomer.
The light bulb lit up that day because of the battery from the electrician according to the customer.
Umilaw ba ang bombilya ayon sa elektrisista? (Did the light bulb light up according to the electrician?)
51. Kumulo' noong isang gabi ang tubig ayon sa estudyante dahil sa kemikal mula sa sayantipiko.
The water boiled one evening according to the student because of the chemical from the scientist.
Kumulo' ba ang tubig dahil sa kemikal? (Did the water boil because of the chemical?)
52. Sumakit noong isang taon ang mga paa ng manlalaro dahil sa paligsahan ayon sa nars.
The feet felt painful last year according to the player because of the game according to the nurse.
Sumakit ba ang mga kamay ayon sa manlalaro? (Did the hands feel painful according to the player?)
53. Bumukas siguro kagabi ang kabinet sa kwarto dahil sa pusa ayon sa babae.
The cabinet in the room perhaps opened up last night because of the cat according to the woman.
Bumukas ba ang kabinet dahil sa pusa? (Did the cabinet open because of the cat?)
54. Nabulok kaninang hapon ang prutas sa kahon mula sa tindera ayon sa mamimili.
The fruits in the box rotted this afternoon from the seller according to the buyer.
Nabulok ba ang gulay sa kahon? (Did the vegetables rot in the box?)
55. Nalanta yata kanina ang halaman mula sa lalaki dahil sa tubig ayon sa hardinera.
The plant from the man perhaps withered yesterday because of the water according to the gardener.
Nalanta ba ang halaman mula sa lalaki? (Did the plant from the man wither?)
56. Napanis kahapon ng gabi ang ulam mula sa kapitbahay dahil sa init ayon sa nanay.
The sidedish from the neighbor spoiled last night because of the heat according to the mother.
Napanis ba ang ulam dahil sa init? (Did the sidedish spoil from the heat?)
57. Natuyo' kaninang tanghali ang labada mula sa dalaga dahil sa hangin ayon sa labandera.
The clothes from the young woman dried up this afternoon because of the wind according to the laundry woman.

Natuyo' ba ang labada dahil sa araw? (Did the clothes dry up because of the sun?)

58. Nag-apoy marahil kanina ang damuhan dahil sa sigarilyo mula sa lalaki ayon sa babae.
The grasses perhaps burned earlier because of the cigarette from the man according to the woman.

Nag-apoy ba ang damuhan dahil sa siga? (Did the grasses burn because of the bonfire?)

59. Nagdilim noong isang araw ang paligid sa pagsabog mula sa suspek ayon sa saksi.
The surrounding darkened the other day because of the explosion from the suspect according to the witness.

Nagdilim ba ang paligid sa simbahan ayon sa saksi? (Did the surrounding darken according to the suspect?)

60. Nagkasya noong isang gabi ang pera mula sa bangko sa dalaga alang-alang sa lalaki.
The money from the bank (was) adequately used one evening by the woman for the man.

Nagkulang ba ang pera mula sa bangko? (Was the money from the bank insufficient?)

Locative Voice

61. Hinigaan kanina lang ng babae ang kama buhat sa kaibigan sa kwarto.
The woman lay down on the bed from the friend just a while ago in the room.
Hinigaan ba ng lolo ang kama? (Did the grandfather lie down on the bed?)

62. Ina:butan kanina umaga ng ulan ang babae sa opisina kasama ng kaibigan.
The woman with the friend caught the rain this morning from the office.
Ina:butan ba ng ulan ang babae? (Did the woman get rained on?)

63. Tinambayan marahil kahapon ng lalaki ang upuan sa simbahan ayon sa pari.
The man perhaps remained on the bench in the church yesterday according to the priest.

Tinambayan ba ang upuan sa simbahan? (Was the chair in the church lingered on (by someone)?)

64. Tinalunan nang araw na iyon ng bata ang trampolin sa palaruan laban sa kapatid.
The child jumped that day on the trampoline in the playground with the sibling.

Inupuan ba ng bata ang trampoline? (Did the child sit on the trampoline?)

65. Tinitigan noong isang buwan ng pulis ang tsuper sa parada dahil sa plaka.
The police stared at the driver in the parking lot last month because of the banner.
Tinitigan ba ng pulis ang tsuper? (Did the police stare at the driver?)

66. Tinangisan noong isang linggo ng lalaki ang sulat mula sa koreo dahil sa babae.
The man wept with the letter last week from the post office because of the woman.

Tinawanan ba ng lalaki ang sulat? (Did the man laugh at the letter?)

67. Hinawakan noong makalawa ng doktor ang kamay hanggang sa braso ayon sa nars.
The doctor touched the hand up to the arm according to the nurse the day before yesterday.

Hinawakan ba ng doktor ang braso? (Did the doctor touch the arm?)

68. Nilayasan siguro kahapon ng matanda ang bahay dahil sa ingay ng bata.
The old man perhaps left the house yesterday because of the noise of the child.

Nilayasan ba ang bahay dahil sa ingay? (Was the house left (by someone) because of the noise?)

69. Dinesisyunan kaninang hapon ng huwes ang kaso sa korte ayon sa mamamahayag.
The judge decided the case this afternoon in the court according to the reporter. 
Dinesisyunan ba ang kaso sa opisina? (Was the case decided in the office?)

70. Pinagtalunan yata kahapon ng mga bata ang laruan mula sa tindahan ayon sa tagapag-alaga.
The children perhaps fought (over) for the toys from the store yesterday according to the caretaker.
Pinagtalunan ba ng mga bata ang pagkain? (Did the children fight over the food?)

71. Pinagtagpuan kahapon ng gabi ng mag-asawa ang gate sa ospital ayon sa nars.
The couple saw the gate in the hospital yesterday evening according to the nurse.
Pinagtagpuan ba ang gate ayon sa nars? (Was the gate a meeting place according to the nurse?)

72. Pinagpahtingahan kaninang tanghali ng magsasaka ang batis sa gubat ayon sa bata.
The farmer rested by the stream this afternoon according to the child.
Pinagpahtingahan ba ang batuhan sa gubat? (Did the farmer rest on the rocks in the forest?)
Appendix E
Stimuli Items and Comprehension Questions in Experiments 3 to 5

Note that these are the approximate English translations of the Tagalog sentences; the order of words in the Tagalog sentence differs from the English translations. The data and other materials will be available in my OSF account: https://osf.io/p37f2/

Critical Items

1. Because of the accident, the two poor families searched for three intelligent doctors in the province earlier.
   AV VAP: Dahil sa aksidente humanap kanina ang dalawang dukhang pamilya ng tatlong matatalinong doktor sa probinsya.
   Humanap ba ang mga pamilya ng mga doktor? (Did the families search for the doctors?)
   AV VPA: Dahil sa aksidente humanap kanina ng tatlong matatalinong doktor ang dalawang dukhang pamilya sa probinsya.
   Humanap ba ng mga doktor ang mga pamilya? (Did the families search for the doctors?)
   PV VAP: Dahil sa aksidente hinanap kanina ang dalawang dukhang pamilya ang tatlong matatalinong doktor sa probinsya.
   Hinanap ba ng mga pamilya ang mga doktor? (Did the families search for the doctors?)
   PV VPA: Dahil sa aksidente hinanap kanina ng dalawang dukhang pamilya ang tatlong matatalinong doktor sa probinsya.
   Hinanap ba ng mga doktor ang mga pamilya? (Did the families search for the doctors?)

2. According to the news, the two gentle priests witnessed (the marriage) of two thoughtful couples at the beach earlier.
   AV VAP: Ayon sa balita nagkasal kanina ng limang maasikasong magkasintahan sa dalampasigan.
   Nagkasal ba ang mga pari ng mga magkasintahan? (Did the priests witness (the marriage) of the couples?)
   AV VPA: Ayon sa balita nagkasal kanina ng limang maasikasong magkasintahan ang dalawang maaamong pari sa dalampasigan.
   Nagkasal ba ng mga magkasintahan ang mga pari? (Did the priests witness (the marriage) of the couples?)
   PV VAP: Ayon sa balita ikinasal kanina ng dalawang maaamong pari ng limang maasikasong magkasintahan sa dalampasigan.
   Ikinasal ba ng mga pari ang mga magkasintahan? (Did the priests witness (the marriage) of the couples?)
   PV VPA: Ayon sa balita ikinasal kanina ng limang maasikasong magkasintahan ng dalawang maaamong pari sa dalampasigan.
   Ikinasal ba ang mga magkasintahan ng mga pari? (Did the priests witness (the marriage) of the couples?)

3. In spite of the trouble, the two rich mothers became concerned (with) the three dirty children about the incident last night.
AV VAP: Sa kabila ng gulo nag-intindi kagabi ang dalawang maririwasang nanay ng tatlong maruruming anak sa pangyayari.
Nag-intindi ba ng mga anak ang mga nanay? (Did the children become concerned with the mothers?)
AV VPA: Sa kabila ng gulo nag-intindi kagabi ng tatlong maruruming anak ang dalawang maririwasang nanay sa pangyayari.
Nag-intindi ba ng mga nanay ang mga anak? (Did the children become concerned with the mothers?)
PV VAP: Sa kabila ng gulo inintindi kagabi ng dalawang maririwasang nanay ang tatlong maruruming anak sa pangyayari.
Inintindi ba ng mga anak ang mga nanay? (Did the children become concerned with the mothers?)
PV VPA: Sa kabila ng gulo inintindi kagabi ang tatlong maruruming anak ng dalawang maririwasang nanay sa pangyayari.
Inintindi ba ang mga nanay ang mga anak? (Did the children become concerned with the mothers?)

4. Despite the fear, the three stout men saved the three crying children from the accident last night.
AV VAP: Sa kabila ng takot nagligtas kagabi ang tatlong matatabang lalaki ng limang iyaking bata mula sa aksidente.
Nagligtas ba ang mga lalaki ng mga bata? (Did the men save the children?)
AV VPA: Sa kabila ng takot nagligtas kagabi ng limang iyaking bata ang tatlong matatabang lalaki mula sa aksidente.
Nagligtas ba ng mga bata ang mga lalaki? (Did the men save the children?)
PV VAP: Sa kabila ng takot iniligtas kagabi ang tatlong matatabang lalaki ang limang iyaking bata mula sa aksidente.
Iniligtas ba ng mga bata ang mga lalaki? (Did the men save the children?)
PV VPA: Sa kabila ng takot iniligtas kagabi ang limang iyaking bata ng tatlong matatabang lalaki mula sa aksidente.
Iniligtas ba ang mga bata ng mga lalaki? (Did the men save the children?)

5. Because of greed, the three evil owners cheated the seven poor customers in the store earlier.
AV VAP: Dahil sa kasakiman nanloko kanina ang tatlong masasamang may-ari ng pitong maralitang kustomer sa tindahan.
Nanloko ba ang mga kustomer ng mga may-ari? (Did the customers cheat the owners?)
AV VPA: Dahil sa kasakiman nanloko kanina ng pitong maralitang kustomer ang tatlong masasamang may-ari sa tindahan.
Nanloko ba ng mga may-ari ang mga kustomer? (Did the customers cheat the owners?)
PV VAP: Dahil sa kasakiman niloko kanina ng tatlong masasamang may-ari ang pitong maralitang kustomer sa tindahan.
Niloko ba ng mga may-ari ang mga kustomer? (Did the customers cheat the owners?)
PV VPA: Dahil sa kasakiman niloko kanina ang pitong maralitang kustomer ng tatlong masasamang may-ari sa tindahan.
Niloko ba ang mga may-ari ng mga kustomer? (Did the customers cheat the owners?)
6. Because of failure, the two rich captains abducted the eight agile staff for the fight yesterday.
   AV VAP: Dahil sa pagkabigo nang-agaw kahapon ang dalawang maririwasang kapitan ng walong maliliksing tauhan para sa laban.
   Nang-agaw ba ang mga kapitan ng mga tauhan? (Did the captains abduct the staff?)
   AV VPA: Dahil sa pagkabigo umagaw kahapon ng walong maliliksing tauhan ang dalawang maririwasang kapitan para sa laban.
   Umagaw ba ng mga tauhan ang mga kapitan? (Did the captains abduct the staff?)
   PV VAP: Dahil sa pagkabigo inagaw kahapon ng dalawang maririwasang kapitan ang walong maliliksing tauhan para sa laban.
   Inagaw ba ng mga tauhan ang mga kapitan? (Did the captains abduct the staff?)

7. According to the story, the two weak pregnant women gave birth to two smiling babies in the hospital yesterday.
   AV VAP: Ayon sa kwento nagsilang kagabi ang dalawang mahihinang buntis ng dalawang palangiting sanggol sa ospital.
   Nagsilang ba ang mga buntis ng mga sanggol? (Did the pregnant women give birth to babies?)
   AV VPA: Ayon sa kwento nagsilang kagabi ng dalawang mahihinang buntis ang dalawang palangiting sanggol sa ospital.
   Nagsilang ba ng mga sanggol ang mga buntis? (Did the pregnant women give birth to babies?)
   PV VAP: Ayon sa kwento inisilang kagabi ng dalawang mahihinang buntis ang dalawang palangiting sanggol sa ospital.
   Isinilang ba ng mga buntis ang mga sanggol? (Did the pregnant women give birth to babies?)
   PV VPA: Ayon sa kwento isinilang kagabi ng dalawang palangiting sanggol ng dalawang mahihinang buntis sa ospital.
   Isinilang ba ang mga sanggol ng mga buntis? (Did the pregnant women give birth to babies?)

8. Because of jealousy, the two snobbish wives pried the two deceitful husbands in the park earlier.
   AV VAP: Dahil sa selos nag-usisa kanina ang dalawang supladang misis ng dalawang bolerong mister sa pasyalan.
   Nag-usisa ba ang mga misis ng mga mister? (Did the wives pry the husbands?)
   AV VPA: Dahil sa selos nag-usisa kanina ng dalawang bolerong mister ang dalawang supladang misis sa pasyalan.
   Nag-usisa ba ng mga mister ang mga misis? (Did the wives pry the husbands?)
   PV VAP: Dahil sa selos inusisa kanina ng dalawang supladang misis ang dalawang bolerong mister sa pasyalan.
   Inusisa ba ng mga misis ang mga mister? (Did the wives pry the husbands?)
   PV VPA: Dahil sa selos inusisa kanina ang dalawang bolerong mister ng dalawang supladang misis sa pasyalan.
Inusisa ba ang mga mister ng mga misis? (Did the wives pry the husbands?)

9. Because of excitement, the five beautiful ladies teased the two good-looking men last night.
   AV VAP: Dahil sa kilig nanukso kagabi ang limang magagandang dalaga ng dalawang gwapong binata sa text.
   Nanukso ba ang mga binata ng mga dalaga? (Did the men tease the ladies?)
   AV VPA: Dahil sa kilig nanukso kagabi ng dalawang gwapong binata ang limang magagandang dalaga sa text.
   Nanukso ba ng mga dalaga ang mga binata? (Did the men tease the ladies?)
   PV VAP: Dahil sa kilig tinukso kagabi ng limang magagandang dalaga ang dalawang gwapong binata sa text.
   Tinukso ba ng mga binata ang mga dalaga? (Did the men tease the ladies?)
   PV VPA: Dahil sa kilig tinukso kagabi ng dalawang gwapong binata ng limang magagandang dalaga sa text.
   Tinukso ba ang mga dalaga ang mga binata? (Did the men tease the ladies?)

10. According to the report, the two former presidents forgave the five fortunate prisoners in the jail room yesterday.
    AV VAP: Ayon sa ulat nagpatawad kahapon ang dalawang dating pangulo ng limang masuswerteng bilanggo sa kulungan.
    Nagpatawad ba ang mga pangulo ng mga bilanggo? (Did the presidents forgive the prisoners?)
    AV VPA: Ayon sa ulat nagpatawad kahapon ng limang masuswerteng bilanggo ang dalawang dating pangulo sa kulungan.
    Nagpatawad ba ng mga bilanggo ang mga pangulo? (Did the presidents forgive the prisoners?)
    PV VAP: Ayon sa ulat pinatawad kahapon ng dalawang dating pangulo ang limang masuswerteng bilanggo sa kulungan.
    Pinatawad ba ng mga pangulo ang mga bilanggo? (Did the presidents forgive the prisoners?)
    PV VPA: Ayon sa ulat pinatawad kahapon ng limang masuswerteng bilanggo ng dalawang dating pangulo sa kulungan.
    Pinatawad ba ang mga bilanggo ng mga pangulo? (Did the presidents forgive the prisoners?)

11. For the activity, the two quick bosses organized the ten new staff in the factory last night.
    AV VAP: Para sa gawain nagtipon kagabi ang dalawang matutuling amo ng sampung bagong manggagawa sa pabrika.
    Nagtipon ba ang mga amo ng mga manggagawa? (Did the bosses organize the staff?)
    AV VPA: Para sa gawain nagtipon kagabi ng sampung bagong manggagawa ang dalawang matutuling amo sa pabrika.
    Nagtipon ba ng mga manggagawa ang mga amo? (Did the bosses organize the staff?)
    PV VAP: Para sa gawain tinipon kagabi ng dalawang matutuling amo ang sampung bagong manggagawa sa pabrika.
    Tinipon ba ng mga amo ang mga manggagawa? (Did the bosses organize the staff?)
    PV VPA: Para sa gawain tinipon kagabi ang sampung bagong manggagawa ng dalawang matutuling amo sa pabrika.
Tinipon ba ang mga manggagawa ng mga amo? (Did the bosses organize the staff?)

12. As a form of assistance, the three caring women fetched the two deaf siblings from the school yesterday.
   AV VAP: Bilang tulong nagsundo kahapon ang tatlong maaasikasong babae ng dalawang binging kapatid sa paaralan.
   Nagsundo ba ang mga bingi ng mga babae? (Did the deaf (siblings) fetch the women?)
   AV VPA: Bilang tulong nagsundo kahapon ng dalawang binging kapatid ang tatlong maaasikasong babae sa paaralan.
   Nagsundo ba ng mga babae ang mga kapatid? (Did the deaf (siblings) fetch the women?)
   PV VAP: Bilang tulong sinundo kahapon ng tatlong maaasikasong babae sa paaralan.
   Sinundo ba ng mga kapatid ang mga babae? (Did the deaf (siblings) fetch the women?)
   PV VPA: Bilang tulong sinundo kahapon ang dalawang binging kapatid ng tatlong maaasikasong babae sa paaralan.
   Sinundo ba ang mga babae ng mga bingi? (Did the deaf (siblings) fetch the women?)

13. For the campaign, the five senior candidates called the two young celebrities on the stage yesterday.
   AV VAP: Para sa kampanya tumawag kagabi ang limang matatandang kandidato ng dalawang batang artista sa entablado.
   Tumawag ba ang mga artista ng mga kandidato? (Did the celebrities call the candidates?)
   AV VPA: Para sa kampanya tumawag kagabi ng dalawang batang artista ang limang matatandang kandidato sa entablado.
   Tumawag ba ng mga kandidato ang mga artista? (Did the celebrities call the candidates?)
   PV VAP: Para sa kampanya tinawag kagabi ng limang matatandang kandidato ang dalawang batang artista sa entablado.
   Tinawag ba ng mga artista ang mga kandidato? (Did the celebrities call the candidates?)
   PV VPA: Para sa kampanya tinawag kagabi ang dalawang batang artista ng limang matatandang kandidato sa entablado.
   Tinawag ba ang mga kandidato ng mga artista? (Did the celebrities call the candidates?)

14. Because of excitement, the five rich singers greeted the eight lucky fans in the show yesterday.
   AV VAP: Dahil sa tuwa bumati kahapon ang limang mayayamang mang-aawit ng walong masuswerteng tagahanga sa labas.
   Bumati ba ang mga tagahanga ng mga mang-aawit? (Did the fans greet the singers?)
   AV VPA: Dahil sa tuwa bumati kahapon ng walong masuswerteng tagahanga ang limang mayayamang mang-aawit sa labas.
   Bumati ba ng mga mang-aawit ang mga tagahanga? (Did the fans greet the singers?)
   PV VAP: Dahil sa tuwa binati kahapon ng limang mayayamang mang-aawit ang walong masuswerteng tagahanga sa labas.
   Binati ba ang mga tagahanga ang mga mang-aawit? (Did the fans greet the singers?)
   PV VPA: Dahil sa tuwa binati kahapon ang walong masuswerteng tagahanga ng limang mayayamang mang-aawit sa labas.
   Binati ba ang mga mang-aawit ng mga tagahanga? (Did the fans greet the singers?)

15. Because of sadness, the two cheerful grandmothers hugged the three tanned grandchildren in the airport earlier.
16. Because of happiness, the three bald priests praised the ten devout children in the church yesterday.

AV VAP: Dahil sa galak pumuri kahapon ang tatlong kalbong pari ng sampung madadasaling bata sa simbahan.

Pumuri ba ang mga pari ng mga bata? (Did the priests praise the children?)

PV VAP: Dahil sa galak pinuri kahapon ang sampung madadasaling bata ng tatlong kalbong pari sa simbahan.

Pinuri ba ng mga pari ang mga bata? (Did the priests praise the children?)

17. Because of the secret, the three tale-teller ladies pulled the two talkative friends near the tree last night.

AV VAP: Dahil sa sikreto nanghila kagabi ang tatlong tsismosang babae ng dalawang intrigerang kaibigan sa puno.

Nanghila ba ang mga babae ng mga kaibigan? (Did the ladies pull the friends?)

PV VAP: Dahil sa sikreto hinila kagabi ng tatlong tsismosang babae ang dalawang intrigerang kaibigan sa puno.

Hinila ba ng mga babae ng mga kaibigan? (Did the ladies pull the friends?)

18. Because of the game, the five meek mayors convened with the ten shy farmers in the farm earlier.

AV VAP: Dahil sa laro nagpulong kanina ang limang maaamong alkalde ng sampung mahiyaing magsasaka sa bukirin.
Nagpulong ba ang mga alkalde ng mga magsasaka? (Did the mayors convene with the farmers?)

AV VPA: Dahil sa laro nagpulong kanina ng sampung mahiyaing magsasaka ang limang maaamong alkalde sa bukirin.

PV VAP: Dahil sa laro pinulong kanina ng limang maaamong alkalde ang sampung mahiyaing magsasaka sa bukirin.

Pinulong ba ang mga magsasaka ng mga alkalde? (Did the mayors convene with the farmers?)

AV VPA: Dahil sa laro nagpulong kanina ng sampung mahiyaing magsasaka ang limang maaamong alkalde sa bukirin.

PV VAP: Dahil sa laro pinulong kanina ang sampung mahiyaing magsasaka ng limang maaamong alkalde sa bukirin.

19. Because of frustration, the two strict teachers hit the three dirty students in the class yesterday.

AV VAP: Dahil sa inis namalo kahapon ng tatlong salaulang estudyante sa klase.

PV VAP: Dahil sa inis pinalo kahapon ng tatlong salaulang estudyante sa klase.

Namalo ba ang mga estudyante ng mga guro? (Did the students hit the teachers?)

PV VAP: Dahil sa inis namalo kahapon ng tatlong salaulang estudyante ang dalawang istriktang guro sa klase.

Namalo ba ng mga guro ang mga estudyante? (Did the students hit the teachers?)

PV VAP: Dahil sa inis pinalo kahapon ng dalawang istriktang guro ang tatlong salaulang estudyante.

Pinalo ba ng mga estudyante ang mga guro? (Did the students hit the teachers?)

PV VAP: Dahil sa inis pinalo kahapon ng dalawang istriktang guro sa klase.

Pinalo ba ang mga guro ng mga estudyante? (Did the students hit the teachers?)

20. For the tournament, the two serious trainers organized the ten new athletes in the park last night.

AV VAP: Para sa paligsahan naghanay kagabi ang dalawang seryosong trainer ng sampung bagong atleta sa parke.

PV VAP: Para sa paligsahan naghanay kagabi ang sampung bagong atleta sa parke.

Naghanay ba ang mga trainer ng mga atleta? (Did the trainers organize the athletes?)

PV VAP: Para sa paligsahan naghanay kagabi ng dalawang seryosong trainer sa parke.

Naghanay ba ng mga atleta ang mga trainer? (Did the trainers organize the athletes?)

PV VAP: Para sa paligsahan hinanay kagabi ng dalawang seryosong trainer ang sampung bagong atleta sa parke.

Hinanay ba ng mga trainer ang mga atleta? (Did the trainers organize the athletes?)

PV VAP: Para sa paligsahan hinanay kagabi ang sampung bagong atleta ng dalawang seryosong trainer sa parke.

Hinanay ba ang mga atleta ng mga trainer? (Did the trainers organize the athletes?)

21. According to the report, the five abusive gentlemen mistreated the three careless guards near the gate last night.
AV VAP: Ayon sa pahayag nang-api kagabi ang limang abusadong ginoo ng tatlong pabayang gwardya sa gate.
Nang-api ba ang mga gwardya ng mga ginoo? (Did the guards mistreat the gentlemen?)
AV VPA: Ayon sa pahayag nang-api kagabi ng tatlong pabayang gwardya ang limang abusadong ginoo sa gate.
Nang-api ba ng mga ginoo ang mga gwardya? (Did the guards mistreat the gentlemen?)
PV VAP: Ayon sa pahayag inapi kagabi ng limang abusadong ginoo ang tatlong pabayang gwardya sa gate.
Inapi ba ng mga gwardya ang mga ginoo? (Did the guards mistreat the gentlemen?)
PV VPA: Ayon sa pahayag inapi kagabi ng limang abusadong ginoo ang tatlong pabayang gwardya sa gate.
Inapi ba ang mga ginoo ng mga gwardya? (Did the guards mistreat the gentlemen?)

22. Because of the incident, the two reliable lawyers prosecuted three vicious culprits at work yesterday.
AV VAP: Dahil sa gulo nag-usig kahapon ang dalawang matitinong abogado ng tatlong salbaheng salarin sa trabaho.
Nag-usig ba ang mga abogado ng mga salarin? (Did the lawyers prosecute the culprits?)
AV VPA: Dahil sa gulo nag-usig kahapon ng tatlong salbaheng salarin ang dalawang matitinong abogado sa trabaho.
Nag-usig ba ng mga salarin ang mga abogado? (Did the lawyers prosecute the culprits?)
PV VAP: Dahil sa gulo inusig kahapon ng dalawang matitinong abogado ang tatlong salbaheng salarin sa trabaho.
Inusig ba ng mga abogado ang mga salarin? (Did the lawyers prosecute the culprits?)
PV VPA: Dahil sa gulo inusig kahapon ang tatlong salbaheng salarin ng dalawang matitinong abogado sa trabaho.
Inusig ba ang mga salarin ng mga abogado? (Did the lawyers prosecute the culprits?)

23. For the school opening, the three busy drivers accommodated the five cheerful passengers in the jeep last night.
AV VAP: Para sa pasukan nag-asikaso kagabi ang tatlong abalang drayber ng limang masasayang pasahero sa jeep.
Nag-asikaso ba ang mga pasahero ng mga drayber? (Did the passengers accommodate the drivers?)
AV VPA: Para sa pasukan nag-asikaso kagabi ng limang masasayang pasahero ang tatlong abalang drayber sa jeep.
Nag-asikaso ba ng mga drayber ang mga pasahero? (Did the passengers accommodate the drivers?)
PV VAP: Para sa pasukan inasikaso kagabi ng tatlong abalang drayber ang limang masasayang pasahero sa jeep.
Inasikaso ba ng mga pasahero ang mga drayber? (Did the passengers accommodate the drivers?)
PV VPA: Para sa pasukan inasikaso kagabi ang limang masasayang pasahero ng tatlong abalang drayber sa jeep.
Inasikaso ba ang mga drayber ng mga pasahero? (Did the passengers accommodate the drivers?)
24. For the consultation, the three multi-racial doctors entertained the five anxious patients from the bed yesterday.

AV VAP: Para sa konsultasyon nang-aliw kahapon ang tatlong mestisong doktor ng limang balisang pasyente mula sa kama.
Nang-aliw ba ang mga pasyente ng mga doktor? (Did the patients entertain the doctors?)
AV VPA: Para sa konsultasyon inaliw kahapon ng tatlong mestisong doktor ang limang balisang pasyente mula sa kama.
Inaliw ba ng mga pasyente ang mga doktor? (Did the patients entertain the doctors?)
PV VPA: Para sa konsultasyon inaliw kahapon ang limang balisang pasyente ng tatlong mestisong doktor mula sa kama.
Inaliw ba ang mga doktor ng mga pasyente? (Did the patients entertain the doctors?)

25. Because of anger, the two judgmental grandfathers drove away the three lame beggars on the street yesterday.

AV VAP: Dahil sa galit nagtaboy kahapon ang dalawang mapangmatang lolo ng tatlong lumpong ulila sa kalsada.
Nagtaboy ba ang mga lolo ng mga ulila? (Did the grandfathers drive away the beggars?)
AV VPA: Dahil sa galit nagtaboy kahapon ang tatlong lumpong ulila ang dalawang mapangmatang lolo sa kalsada.
Nagtaboy ba ng mga ulila ang mga lolo? (Did the grandfathers drive away the beggars?)
PV VPA: Dahil sa galit nagtaboy kahapon ang tatlong lumpong ulila ng dalawang mapangmatang lolo sa kalsada.
Tinaboy ba ang mga lolo ng mga ulila? (Did the grandfathers drive away the beggars?)
PV VPA: Dahil sa galit nagtaboy kahapon ang tatlong lumpong ulila ng dalawang mapangmatang lolo sa kalsada.
Tinaboy ba ng mga ulila ang mga lolo? (Did the grandfathers drive away the beggars?)

26. Because of love, the two rich ladies eloped with the two stocky men from the restaurant yesterday.

AV VAP: Dahil sa pagmamahal nagtanan kahapon ang dalawang mayayamang dalaga ng dalawang matitipunong binata mula sa restoran.
Nagtanan ba ang mga binata ng mga dalaga? (Did the men elope with the ladies?)
AV VPA: Dahil sa pagmamahal nagtanan kahapon ng dalawang matitipunong binata ang dalawang mayayamang dalaga mula sa restoran.
Nagtanan ba ng mga dalaga ang mga binata? (Did the men elope with the ladies?)
PV VPA: Dahil sa pagmamahal tinaan kahapon ng dalawang mayayamang dalaga ang dalawang matitipunong binata mula sa restoran.
Tinanana ba ng mga binata ang mga dalaga? (Did the men elope with the ladies?)
PV VPA: Dahil sa pagmamahal tinaan kahapon ang dalawang matitipunong binata ng dalawang mayayamang dalaga mula sa restoran.
Tinanana ba ang mga dalaga ng mga binata? (Did the men elope with the ladies?)

27. For the exercise, the five hardworking nurses assisted the ten weak patients from the bed last night.
AV VAP: Para sa eholdersy nag-ayak kagabi ang limang matitiyagang nars ng sampung mahihinang pasyente mula sa kama.
Nag-ayak ba ang mga nars ng mga pasyente? (Did the nurses assist the patients?)
AV VPA: Para sa eholdersy nag-ayak kagabi ng sampung mahihinang pasyente ang limang matitiyagang nars mula sa kama.
Nag-ayak ba ng mga pasyente ang mga nars? (Did the nurses assist the patients?)
PV VAP: Para sa eholdersy inakay kagabi ng limang matitiyagang nars ang sampung mahihinang pasyente mula sa kama.
Inakay ba ng mga nars ang mga pasyente? (Did the nurses assist the patients?)
PV VPA: Para sa eholdersy inakay kagabi ang sampung mahihinang pasyente ng limang matitiyagang nars mula sa kama.
Inakay ba ang mga pasyente ng mga nars? (Did the nurses assist the patients?)

28. According to the report, the three convinced judges investigated the seven negligent senators in the court last night.
AV VAP: Ayon sa ulat nagsiyasat kagabi ang tatlong kumbinsidong huwes ng pitong pabayang senator sa korte.
Nagsiyasat ba ang mga senador ng mga huwes? (Did the senators investigate the judges?)
AV VPA: Ayon sa ulat nagsiyasat kagabi ng pitong pabayang senator ang tatlong kumbinsidong huwes sa korte.
Nagsiyasat ba ng mga huwes ang mga senador? (Did the senators investigate the judges?)
PV VAP: Ayon sa ulat siniyasat kagabi ng tatlong kumbinsidong huwes ang pitong pabayang senador sa korte.
Siniyasat ba ng mga senador ang mga huwes? (Did the senators investigate the judges?)
PV VPA: Ayon sa ulat siniyasat kagabi ang pitong pabayang senador ng tatlong kumbinsidong huwes sa korte.
Siniyasat ba ang mga huwes ng mga senador? (Did the senators investigate the judges?)

29. According to the news, the two selfish presidents cheated the five wise rivals in the election yesterday.
AV VAP: Ayon sa balita nandaya kahapon ang dalawang gahamang presidente ng limang matatalinong katunggali sa eleksyon.
Nandaya ba ang mga presidente ng mga katunggali? (Did the presidents cheat the rivals?)
AV VPA: Ayon sa balita nandaya kahapon ng limang matatalinong katunggali ang dalawang gahamang presidente sa eleksyon.
Nandaya ba ng mga katunggali ang mga presidente? (Did the presidents cheat the rivals?)
PV VAP: Ayon sa balita dinaya kahapon ng dalawang gahamang presidente ang limang matatalinong katunggali sa eleksyon.
Dinaya ba ng mga presidente ang mga katunggali? (Did the presidents cheat the rivals?)
PV VPA: Ayon sa balita dinaya kahapon ang limang matatalinong katunggali ng dalawang gahamang presidente sa eleksyon.
Dinaya ba ng mga katunggali ng mga presidente? (Did the presidents cheat the rivals?)

30. According to the story, the two forgetful seniors complained (of) the eight fake policemen in the station last night.
AV VAP: Ayon sa kwesto nagreklamo kagabi ang dalawang ulyang matanda ng walong pekeng pulis sa himpilan.
Nagreklamo ang mga pulis ng mga matanda? (Did the policemen complain (of) the seniors?)
AV VPA: Ayon sa kwento nagreklamo kagabi ng walang pekeng pulis ang dalawang ulyaning matanda sa himpilan.
Nagreklamo ba ng mga matanda ang mga pulis? (Did the policemen complain (of) the seniors?)
PV VAP: Ayon sa kwento inireklamo kagabi ng dalawang ulyaning matanda ang walang pekeng pulis sa himpilan.
Inireklamo ba ng mga pulis ang mga matanda? (Did the policemen complain (of) the seniors?)
PV VPA: Ayon sa kwento inireklamo kagabi ng dalawang ulyaning matanda ang walang pekeng pulis sa himpilan.
Inireklamo ba ang mga matanda ng mga pulis? (Did the policemen complain (of) the seniors?)

31. Out of anger the three poor personnel stabbed the five wealthy enemies in the crime yesterday.
AV VAP: Dahil sa galit nanaksak ak kahapon ang tatlong pobreng tauhan ng limang mayayamang kaaway sa krimen.
Nanaksak ba ang mga kaaway ng mga tauhan? (Did the enemies stab the personnel?)
AV VPA: Dahil sa galit nanaksak kahapon ng limang mayayamang kaaway ang tatlong pobreng tauhan sa krimen.
Nanaksak ba ng mga tauhan ang mga kaaway? (Did the enemies stab the personnel?)
PV VAP: Dahil sa galit sinaksak kahapon ng limang mayayamang kaaway ang tatlong pobreng tauhan sa krimen.
Sinaksak ba ng mga kaaway ang mga tauhan? (Did the enemies stab the personnel?)
PV VPA: Dahil sa galit sinaksak kahapon ng limang mayayamang kaaway ng tatlong pobreng tauhan sa krimen.
Sinaksak ba ang mga tauhan ng mga kaaway? (Did the enemies stab the personnel?)

32. According to the report, the seven prudent guards attacked the two cunning thieves in town yesterday.
AV VAP: Ayon sa ulat sumalakay kahapon ang pitong siguristang bantay ng dalawang tusong magnanakaw sa bayan.
Sumalakay ba ang mga magnanakaw ng mga bantay? (Did the thieves attack the guards?)
AV VPA: Ayon sa ulat sumalakay kahapon ng dalawang tusong magnanakaw ang pitong siguristang bantay sa bayan.
Sumalakay ba ng mga bantay ang mga magnanakaw? (Did the thieves attack the guards?)
PV VAP: Ayon sa ulat sinalakay kahapon ng pitong siguristang bantay ang dalawang tusong magnanakaw sa bayan.
Sinalakay ba ng mga magnanakaw ang mga bantay? (Did the thieves attack the guards?)
PV VPA: Ayon sa ulat sinalakay kahapon ang dalawang tusong magnanakaw ng pitong siguristang bantay sa bayan.
Sinalakay ba ang mga bantay ng mga magnanakaw? (Did the thieves attack the guards?)

33. According to the rumor, the three extravagant foreigners invited the ten rich relatives in the country last night.
34. For the banquet, the two hardworking mothers looked for the three languid maids in the market yesterday.

AV VAP: Para sa handaan naghagilap kahapon ang dalawang matitiyagang ina ng tatlong malalamyang kasambahay sa palengke.

Naghagilap ba ang mga kasambahay ng mga ina? (Did the maids look for the mothers?)

AV VPA: Para sa handaan naghagilap kahapon ng tatlong malalamyang kasambahay ang dalawang matitiyagang ina sa palengke.

Naghagilap ba ng mga ina ang mga kasambahay? (Did the maids look for the mothers?)

PV VAP: Para sa handaan hinagilap kahapon ng dalawang matitiyagang ina ang tatlong malalamyang kasambahay sa palengke.

Hinagilap ba ng mga kasambahay ang mga ina? (Did the maids look for the mothers?)

PV VPA: Para sa handaan hinagilap kahapon ng dalawang matitiyagang ina sa palengke.

Hinagilap ba ang mga ina ng mga kasambahay? (Did the maids look for the mothers?)

35. Following the operation, the five quick policemen arrested the three deceitful kidnappers at the crossing last night.

AV VAP: Alinsunod sa operasyon nagdakip kagabi ang limang matutuling pulis ng tatlong madadayang kidnaper sa tawiran.

Nagdakip ba ang mga pulis ng mga kidnaper? (Did the policemen arrest the kidnappers?)

AV VPA: Alinsunod sa operasyon nagdakip kagabi ng tatlong madadayang kidnaper ang limang matutuling pulis sa tawiran.

Nagdakip ba ng mga kidnaper ang mga pulis? (Did the policemen arrest the kidnappers?)

PV VAP: Alinsunod sa operasyon dinakip kagabi ng limang matutuling pulis ang tatlong madadayang kidnaper sa tawiran.

Dinakip ba ng mga pulis ang mga kidnaper? (Did the policemen arrest the kidnappers?)

PV VPA: Alinsunod sa operasyon dinakip kagabi ang tatlong madadayang kidnaper ng limang matutuling pulis sa tawiran.

Dinakip ba ang mga kidnaper ng mga pulis? (Did the policemen arrest the kidnappers?)
36. Because of annoyance, the three wicked sisters kicked the five feeble siblings in the playground earlier.

AV VAP: Dahil sa kulit nanipa kanina ang tatlong salbaheng ate ng limang mahihinang kapatid sa palaruan.
Nanipa ba ang mga ate ng mga kapatid? (Did the sisters kick the siblings?)
AV VPA: Dahil sa kulit nanipa kanina ng limang mahihinang kapatid ang tatlong salbaheng ate sa palaruan.
Nanipa ba ng mga kapatid ang mga ate? (Did the sisters kick the siblings?)
PV VAP: Dahil sa kulit sinipa kanina ang tatlong salbaheng ate sa palaruan.
Sinipa ba ng mga kapatid ang mga ate? (Did the sisters kick the siblings?)
PV VPA: Dahil sa kulit sinipa kanina ng tatlong salbaheng kapatid ng limang mahihinang kapatid sa palaruan.
Sinipa ba ng mga kapatid ng mga ate? (Did the sisters kick the siblings?)

Filler Items (While all alternative pronouns are offered in the English translations, there is only one third person singular pronoun in Tagalog, which was used in the stimuli items.)
1. Naglabas kahapon ng sama ng loob ang dalawang matanda sa madre tungkol sa tatlong kuripot na biyenan na kasama nila.
The two seniors expressed their frustration to the nun about the three stingy father-in-laws who was their companion.
Naglabas ba ng sama ng loob ang mga matanda? (Did the seniors express frustration?)

2. Nagbalak lumapit nang matulin ang tatlong pusa sa dalawang daga sa may mesa sa kusina kaninang tanghali.
The three cats planned to quickly approach the two rats at the kitchen table this afternoon.
Nagbalak bang lumapit ang mga pusa sa mga daga? (Did the cats plan to approach the rats?)

3. Noong isang buwan dumaan nang mabilis ang desperadong bilanggo sa pasilyo na may bitbit sa kamay na dokumento.
Last month, the desperate prisoner quickly walked through the hallway while carrying some documents.
Mabilis bang dumaan ang bilanggo sa pasilyo? (Did the prisoner walk through the hallway quickly?)

4. Umiyak kahapon ang nababahalang dalaga sa limang madre sa kumbento dahil sa kasalanan nila sa kababayan mula sa probinsya.
The anxious lady cried to the five nuns in the convent because of their fault against a fellow mate from the province.
Umiwas ba ang dalaga sa dalawang madre? (Did the lady avoid the two nuns?)

5. Dahil sa sugat sa magkabilang dalat sa limang madre sa klinikong ng doktor.
Because of the wound on both feet caused by the firecracker, the five kids loudly shouted in the doctor’s clinic.
Nasa magkabilang kamay ba ang sugat? (Were the wounds on both hands?)

6. Ninais sumabay ng apat na bata sa nagmamadaling kapatid sa pagsakay ng jeep para makapason sa paaralan.
The four children wanted to accompany the sibling who was rushing to ride the jeep to get to school.

Ninai bang sumabay ng tatlong bata sa jeep? (Did the three children want to ride the jeep?)

   One day, the farmer saw the white airplane fly from Manila to the US despite the typhoon.
   Lumipad ba ang eroplano patungong US? (Did the airplane fly to the US?)

   Because of the therapist, the patient successfully walked from the room to the restroom without assistance.
   Humakbang ba ang pasyente nang walang alalay? (Did the patient walk without assistance?)

   The hardworking baker quickly took a glance at the servant because he/she/they thought he/she/they was/were stealing bread.
   Bumaling ba ng tingin ang panadero? (Did the baker take a glance [at the servant]?)

10. Naisipang bumawi ng estudyante sa mahirap na exam ng mabait na guro para sa grado nito sa klase ngayong taon.
    The student thought of making up from the nice teacher’s difficult exam for his/her/their grade in the class this year.
    Madali ba ang exam ng guro? (Was the teacher’s exam difficult?)

    Last night, the cheerful announcer decided to give a generous prize for the participants in the competition this month.
    Nagpasya ba ang announcer na magbigay ng papremyo? (Did the announcer decide to give a prize?)

12. Nagsadya sa ospital noong isang araw ang nababahalang madre sa mahinhing cashier para tulungan silang magpatingin sa espesyalista.
    The concerned nun went to the hospital one day [to meet] the modest cashier to help her see a specialist.
    Nagsadya ba sa madre ang nababahalang cashier? (Did the cashier go to meet the nun?)

13. Dahil sa pabango madaling naghinala kahapon ng gabi ang morenong lalaki sa sekretarya na may kabit ito sa opisina.
    Because of the perfume, the tanned man quickly suspected the secretary of an affair in the office last night.
    Naghinala ba sa lalaki ang sekretarya? (Did the man suspect the secretary?)

    Because of the feast the other day, the deceitful queen angrily spoke with the king about the gathering in the town.
Nagwika ba sa reyna ang hari? (Did the queen speak with the king?)

15. Habang punung-puno ng galak tinapos ng inhinyero noong isang buwan ang mababang bahay na surpresa para sa may-ari nito.
   With much excitement, the engineer completed the house as a surprise for its owner last month.
   Tinapos ba ng inhinyero ang bahay? (Did the engineer complete the house?)

16. Para gumaling inaral nang maigi ng pamosong doktor ang sakit ng pasyente na wala pang lunas sa ngayon.
   The famous doctor carefully studied the patient’s cureless illness for his/her/their recovery.
   Pamoso ba ang pasyente sa sakit nito? (Was the doctor famous for his/her/their illness?)

17. Para makatipid sa konsumo ng kuryente nagpatay agad ng maliwanag na ilaw ang masinop na binata.
   In order to save electricity, the neat man immediately turned off the bright light.
   Maliwanag ba ang ilaw? (Was the light bright?)

   The tanned man immediately started the game without a prize for the child in the plaza yesterday noon.
   Sinimulan ba ng babae ang palaro? (Did the woman start the game?)

19. Patagong nag-uwi para sa kampanya ang matabang Amerikanong dayuhan ng kontraktor na mayaman sa bansa.
   The stout American foreigner secretly returned home for the rich contractor’s campaign in the country.
   Nag-uwi ba ng kontraktor ang payat na dayuhan? (Did the contractor bring back the thin foreigner?)

20. Pinilit magsama ng ina ng kasambahay sa palengke para tulungan siyang magbuhat ng kasangkapan.
   The mother stubbornly brought the maid to the market to help her carry the materials.
   Nagsama ba ang ina ng kasambahay? (Did the mother bring the maid?)

21. Ang pag-inom ng dalaga ng gamot sa ospital ang nakasama sa kanyang mahinang kalusugan.
   The woman’s intake of the medicine in the hospital worsened her weak health.
   Mahina ba ang kalusugan ng dalaga? (Was the woman’s health weak?)

22. Sinabi ng madre na maimili siya ng gamit sa mall kaysa sa palengke kahit na mahal ang mga ito.
   The nun said that she will buy materials in the mall than in the market despite the materials’ expensive price.
   Maimili ba ang madre ng gamot? (Will the nun be buying medicine?)

23. Ipinagkaila ng pulis na kinontak siya ng ahensya tungkol sa karumal-dumal na krimen kagabi sa lungsod ng Pasay.
   The police denied that he/she/they was/were contacted by the agency about the heinous crime last night in Pasay City.
   Kinontak ba umano na ahente ang pulis? (Did the agent reportedly contact the police?)
24. Dapat kunin ng bawat mag-aaral ang klase ng propesor na iyon dahil kabisado niya ang libro tungkol sa calculus.
   Every student should take that professor’s class because he has the calculus book memorized.
   Kabisado ba ng propesor ang libro sa calculus? (Does the professor have the calculus book memorized?)

25. Maaaring tamaan ng kidlat ang nagkukumpuni ng kuryente sa poste kung gagawin niya ito nang umuulan.
   Lightning can hit the electrician working on the electric post if he/she/they work on it while it is raining.
   May nagkukumpuni ba ng poste ng kuryente? (Was there someone working on the electric post?)

   The Japanese can participate in the trip to Manila for the business that they will begin.
   Sinimulan ba ng mga Hapon ang negosyo? (Did the Japanese begin their business?)

27. Dahil sa takot hindi tumawid ang lalaki sa pedestrian at nag-antay ito ng kasabay sa pagtawid.
   Out of fear, the man did not cross the pedestrian crossing and waited for another person who will cross the street.
   Tumakbo ba ang lalaki sa pedestrian? (Did the man run to cross the pedestrian crossing?)

28. Gustong isumpa noong isang gabi ng mangkukulam ang sinungaling sa gubat dahil sa pagnakaw ng kayamanan.
   The witch wanted to curse the liar in the forest for stealing the treasure.
   Gusto bang isumpa ng mangkukulam ang sinungaling? (Did the witch want to curse the liar?)

29. Noong bakasyon nagpasyal kanina ang piloto ng bisita sa isla gamit ang jeep na de-solar.
   Last vacation, the pilot toured the visitor in the island using the solar-powered jeep.
   Nagpasyal ba ng piloto ang bisita? (Did the pilot tour the visitor?)

30. Noong pista inibig agad-agad ng prinsipe ang dalaga dahil sa husay nito sa pag-awit ng awiting bayan.
   In the last feast, the prince immediately fell for the lady because of her rendition of the folk song.
   Inibig ba ng prinsipe ang dalaga? (Did the prince fall for the lady?)

   In the party, the lady immediately chose a partner for the game held in the community.
   Pumili ba ang dalaga ng kapareha? (Did the lady choose the partner?)

32. Sa kabila ng gulo hinabol nang mabilis ng tagahanga ang mang-aawit sa concert kahit nahihiya siya sa kanya.
   Despite the commotion, the fan quickly chased the singer in the concert despite his/her/their embarrassment for him/her/them.
   Hinabla ba ng tagahanga ang mang-aawit? (Did the fan sue the singer?)

33. Kahit pagod binasa kahapon ng hardinero ang hampaslupa sa hardin gamit ang hose na luma.
Despite exhaustion, the gardener sprayed (water) on the beggar in the garden using the old hose.

Binasa ba ng hardinero ang hampaslupa? (Did the gardener spray [water on] the beggar?)

34. Dahil sa away madiing nagagat ng kalaro ang bata sa palaruan na napansin ng magulang.
   Because of the fight, the playmate bit the child in the playground, which the parent noticed.
   Nangagat ba ang kalaro ng bata? (Did the playmate bite the child?)

35. Dahil sa inis nagyaya nang mabilis ng katulong ang bisita sa daanan upang magpatulong sa gawain.
   Because of annoyance, the maid immediately asked the visitor on the hallway to help in the activity.
   Nagyaya ba ng katulong angbisita? (Did the visitor ask for the maid?)

36. Dahil sa buwisit minura nang malakas ang bilanggo ng pulis sa kulungan kahi ng patong na ang ronda nito.
   Because of frustration, the prisoner loudly sweared at the police in the prison even though his/her/their patrol was over.
   Minura ba nang malakas ang bilanggo ng pulis? (Did the prisoner loudly swear at the police?)

37. Noong pista binugbog nang lubos ang lasing ng kapitbahay dahil sa intriga na kinalat nito sa barangay.
   During the festival, the neighbor severely hit the drunkard because of the gossip he/she/they was/were spreading in the community.
   Binugbog ba ng lasang ang kapitbahay? (Did the drunkard hit the neighbor?)

38. Nahuling nagrekord ang mga mag-aaral ng kanilang sarili para sa ensayo na gagana sa entablado bukas ng umaga.
   The students were late in recording themselves for the stage rehearsals tomorrow morning.
   Nagrekord ba ang mga mag-aaral ng kanilang sarili? (Did the students record themselves?)

39. Inamoy ang sarili ng batang matagal nang hindi naliligo dahil iniwasan siya ng kanyang mga kasamang bata sa palaruan.
   The unsanitary child smelled himself/herself/themselves because he/she/they was/were being avoided by his playmates in the playground.
   Inamoy ba ang sarili ng batang hindi naliligo? (Did the unsanitary child smell himself/herself/themselves?)

40. Madiing nagkamot ng sarili niya ang lalaki sa kanyang paa dahil sa kagat ng langgam galing sa punso.
   The man scratched his own foot because he was bitten by the ant from the mound.
   Nagkamot ba ng sarili niya ang lalaki? (Did the man scratch himself?)

41. Madiing sinabon ang sarili niya ng batang marumi pagkatapos maglaro sa hardin kung saan kasama niya ang kaklase.
   The dirty child washed himself/herself/themselves after playing in the garden with his/her/their playmate.
Sinabon ba ang sarili niya ng batang marumi? (Did the child wash himself/herself/themselves?)

42. Binalot nang mahigpit ng maysakit ang sarili niya dahil sa lamig na dulot ng airon sa ospital.
   The patient tightly wrapped himself/herself/themselves because of the cold temperature in the hospital.
   Binalot ba ng maysakit ang sarili niya? (Did the patient wrap himself/herself/themselves?)

43. Galit na namalo ng bunsong kapatid ang ate dahil ayaw nitong sumunod sa pinag-uutos niya kahapon.
   The sister angrily slapped the younger brother because she did not follow her yesterday.
   Namalo ba ang bunso ng ate? (Did the younger brother slap the sister?)

44. Masayang nangiliti ang tindera ng kanyang mamimili dahil nagkakatuwaan sila sa bagong tsismis ukol sa kanilang kapitbahay.
   The seller cheerfully tickled the customer as they shared gossips about their neighbor.
   Naiinis bang nangiliti ang tindera ng mamimili? (Did the seller angrily tickle the customer?)

45. Galanteng nagbihis ang babae ng magandang damit para dumalo sa isang marangyang pagtitipon.
   The woman elegantly dressed herself in beautiful clothes as she attends a luxurious gathering.
   Galante bang nagbihis ang babae? (Did the woman dress elegantly?)

46. Nagbasang lahat ng mga mag-aaral sa loob ng silid-aklatan noong nakaraang linggo.
   The students all read the thick book in the library last week.
   Makapal ba ang binasang aklat ng mga mag-aaral? (Did the students read the thick book?)

47. Itinabing lahat ng mga mag-aaral ang mga kasangkapan sa isang kabinet na gagamitin sa pagluluto sa susunod na araw.
   The students stored all the utensils in a cabinet for cooking the following day.
   Itinabi bang lahat ng mga mag-aaral ang mga kasangkapan? (Did the students store all the utensils?)

48. Sinirang lahat ng mga lalaki ang mga plakard laban sa pamahalaan habang naglalakad sila sa kalsada kahapon ng tanghali.
   The men destroyed all the placards against the government as they walked on the road yesterday noon.
   Wiinasak ba ng mga lalaki ang mga plakard? (Did the men destroy the placards?)

49. Hinating lahat ng guro ang mga regalo mula sa kanyang mga mag-aaral para ibahagi sa mga mahihirap.
   The teacher divided all the gifts from his/her/their students to be given to the needy.
   Hinati ba ng guro ang relyeno? (Did the teacher divide the relyeno [a kind of fish dish]?)

50. Hiniwang lahat ng nanay ang mga gulay na inani ng tatay mula sa bukid sa Baguio.
   The mother sliced all the vegetables harvested by the father from a farm in Baguio.
   Hiniwa ba ng nanay ang mga prutas? (Did the mother slice the fruits?)
51. Nagtanggal na lahat ng kanilang mga sumbrero ang mga kawal sa isang pagtitipong militar sa parke.
The soldiers all removed their hats in a military gathering in the park.
Nagtanggal ba ng sumbrero ang mga kawal? (Did the soldiers remove their hats?)

52. Nagtunaw na lahat ang mga doktor ng gamot sa tubig na siyang gagamitin sa gamutan sa susunod na araw.
The doctors all dissolved the medicine in the water to be used for treatment the next day.
Nagtunaw ba ang doktora ng gamot? (Did the female doctor dissolve the medicine?)

53. Naghanap na lahat ang mga bata ng kanilang mga tsinelas habang naglaro ng taguan sa palaruan.
The children all searched for their slippers as they played hide and seek in the playground.
Naghanda ba ang mga bata ng kanilang tinselas? (Did the children prepare their slippers?)

54. Naglutong lahat ang mga masisipag na magsasaka na pagkain para sa kaarawan ng kanilang mga asawa.
The hardworking farmers all cooked food for their wives’ birthdays.
Nagluto ba ang mga tamad na magsasaka para sa mga asawa? (Did the lazy farmers cook for their wives?)

55. Para sa tatlong mahirap na ale ipinambitbit kagabi ng lalaki ang pitong mabigat na kariton ng gulay
For the three poor female seniors, the man carried the seven heavy wagons with vegetables last night.
Ipinambitbit ba ng lalaki ang tatlong kariton? (Did the man carry three wagons?)

56. Ipinangpasa ng anak ng tigang na bukid ang asarol ng tigang na bukid para sa negosyanteng nagsisimula pa lang.
Last night, the hardworking and obedient farmers plowed the barren farm with the hoe for the beginning businessman.
Ipinangararo ba ng magsasaka ang asarol? (Did the farmer plow with the hoe?)

57. Ipinang-ani kanina ng binata ang mahabang sungkit ng lolo ng manggang hilaw bilang panregalo sa kaibigan nito.
Earlier, the young man harvested unripe mangoes with the grandfather’s pole as a gift to a friend.
Ipinang-ani ba ang sungkit ng lola? (Was the grandmother’s pole used for harvesting?)

58. Ipinang-araro kababaihan ng masikap at masunuring magsasaka ang asarol ng tigang na bukid para sa negosyanteng nagsisimula pa lang.
Last night, the hardworking and obedient farmers plowed the barren farm with the hoe for the beginning businessman.
Ipinang-araro ba ng magsasaka ang asarol? (Did the farmer plow with the hoe?)

59. Ipinansaway kagabi ng pulis ang pito ang pito na humaharurot na sasakyang maganda na sasakyang maganda para hindi masaktan ang tumawid.
Last night, the police reprimanded the speeding beautiful vehicle with the whistle to avoid harming the pedestrian.
Ipinansaway ba ng pulis ang pito? (Did the police reprimand with the whistle?)
60. Ipinagbalot kahapon ng masipag na nanay ang bata ng aklat sa Sibika gamit ang plastic cover.
   Yesterday, the industrious mother wrapped the child’s book on Civics using the plastic cover.
   Ipinagbalot ba ng nanay ang bata? (Did the mother wrap for the child?)
61. Ipinag-ahit kanina ng binata ang lolong kuba ng puti nitong bigote gamit ang labaha na nabili sa palengke.
   Earlier, the young man shaved the hunchbacked grandfather’s white mustache with the razor bought from the market.
   Ipinag-ahit ba ang binata ng lolo? (Did the young man shave the grandfather?)
62. Ipinagbasag kanina ng yelo ng tindera ng sorbetes ang anak gamit ang ice pick na mapurol.
   The ice cream seller broke the ice for the child with the dull ice pick.
   Ipinagbasag ba ng relo ang anak? (Was the watch broken for the child?)
63. Ipinagdurog siguro kahapon ng nanay ang anak na iyakin ng karot na pinakuluan gamit ang kutsara.
   Yesterday, the mother perhaps mashed the boiled carrots for the crying child with the spoon.
   Ipinagdurog ba ng tatay ang tahimik na anak? (Did the father mash for the quiet child?)
64. Ipinaggupit kanina ng apat na nars ang pasyente ng benda gamit ang guting para sa sugat nito na nagdurugo.
   The four nurses cut the bandage for the patient using the scissors for the bleeding wound.
   Ipinanggupit ba ng apat na nars ang pasyente? (Did the four nurses cut for the patient?)
65. Umulan nang malakas sa siyudad matapos umihip ang malakas at malamig na hangin sa nayon.
   It rained heavily in the city after the strong and cold winds blew through the village.
   Mahina ba ang pag-ulan sa siyudad? (Were there light rains in the city?)
66. Siguradong sumama noong isang linggo ang ginoo sa tatlong mayabang na kumpare mula sa opisina hanggang sa restoran.
   The gentleman certainly joined the three boastful companions from the office to the restaurant last week.
   Sumama ba ang ginang sa mga kumpare? (Did the wives join the companions?)
67. Madaling hiningal kanina Lang umaga sa lahat ng mahabang pagtakbo mula sa baybayin.
   The two industrious men easily gasped for breath this morning from their long jog from the beach.
   Hiningal ba ang tatlong lalaki sa pagtakbo? (Did the three men gasp for breath from their run?)
68. Kani-kanina lang nagkunwari at mahusay na gumanap ang artista bilang clown sa circus alang-alang sa manonood.
   Earlier the actor pretended and performed well as a clown in the circus for the audience.
   Gumanap ba ang artista sa circus? (Did the actor perform in the circus?)
69. Kaninang umaga biglang naglaho ang dalawang matabang kuneho sa kahon dahil sa husay ng mahikero.
This morning the two fat rabbits suddenly disappeared from the box because of the magician’s skill.
Naglaho ba ang dalawang kuneho dahil sa mahikero? (Did the two rabbits disappear because of the magician?)

70. Mariing nagsalita kaninang alas-onse ang matiyaga at bibong guro sa silid-aralan tungkol sa mga libro alang-alang sa estudyante.
At 11’o clock, the hardworking and active teacher emphatically spoke in the classroom regarding the books for the students.
Nagsalita ba ang guro sa silid-aklata? (Did the teacher speak in the library?)

71. Noong madaling araw galit na nagwala ang motorista sa makipot na kalsada dahil sa drayber sa kabila ng trapiko.
Earlier at dawn, the motorist angrily caused a ruckus on the narrow road because of the driver despite the traffic congestion.
Nagwala ba dahil sa motorista ang drayber? (Did the driver cause a ruckus because of the motorist?)

72. Noong makalawa naghirap na lubos ang aktor dahil sa utang mula sa kasino ayon sa asawa nitong sugarol.
Two days ago, the actor became financially impoverished because of the debt from the casino according to his gambler wife.
Yumaman ba ang aktor sa kasino? (Did the actor become rich in the casino?)
Appendix F
Statistical Results of the Reaction Times in Experiment 5 (in Isolation)

Table 18. Results of the mixed effects linear regression models that tested the reaction time differences between conditions per sentence region in Experiment 5. The critical regions are shaded.34

<table>
<thead>
<tr>
<th>Region</th>
<th>$\beta$</th>
<th>SE</th>
<th>$t$</th>
<th>$p$</th>
<th>$\beta$</th>
<th>SE</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Region 1</td>
<td></td>
<td></td>
<td></td>
<td>Region 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>6.97</td>
<td>0.04</td>
<td>178.26</td>
<td>&lt; .001</td>
<td>7.02</td>
<td>0.03</td>
<td>203.22</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Voice</td>
<td>0.03</td>
<td>0.03</td>
<td>1.10</td>
<td>0.28</td>
<td>-0.07</td>
<td>0.03</td>
<td>-2.53</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>WordOrder</td>
<td>0.04</td>
<td>0.02</td>
<td>1.59</td>
<td>0.12</td>
<td>0.004</td>
<td>0.02</td>
<td>0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>Voice:WordOrder</td>
<td>0.001</td>
<td>0.05</td>
<td>0.03</td>
<td>0.98</td>
<td>0.05</td>
<td>0.04</td>
<td>1.34</td>
<td>0.18</td>
</tr>
</tbody>
</table>

| Region 3 |     |       |        | Region 4 |     |       |        |
| Intercept | 6.57 | 0.02 | 294.19 | < .001 | 6.99 | 0.03 | 219.74 | < .001 |
| Voice   | 0.02 | 0.02 | 1.29  | 0.20   | -0.02 | 0.02 | -0.67  | 0.51   |
| WordOrder | -0.01 | 0.02 | -0.37 | 0.71   | 0.03  | 0.04 | 0.84   | 0.41   |
| Voice:WordOrder | 0.05 | 0.04 | 1.27  | 0.21   | -0.04 | 0.04 | -0.81  | 0.43   |

| Region 5 |     |       |        | Region 6 |     |       |        |
| Intercept | 6.53 | 0.02 | 345.96 | < .001 | 6.91 | 0.03 | 212.32 | < .001 |
| Voice   | 0.01 | 0.02 | 0.64  | 0.53   | -0.01 | 0.02 | -0.33  | 0.74   |
| WordOrder | 0.003 | 0.02 | 0.17  | 0.87   | -0.06 | 0.03 | -1.96  | 0.06   |
| Voice:WordOrder | 0.08 | 0.03 | 2.73  | < .01  | -0.004 | 0.05 | -0.08  | 0.93   |

| Region 7 |     |       |        |
| Intercept | 6.50 | 0.02 | 274.60 | < .001 |
| Voice   | 0.002 | 0.02 | 0.12   | 0.91   |
| WordOrder | -0.004 | 0.02 | -0.19  | 0.85   |
| Voice:WordOrder | 0.07 | 0.03 | 2.28  | < .05  |

34Mixed effects linear regression models used
    lmer(logR1 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))
lmer(logR2 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))
lmer(logR3 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))
lmer(logR4 ~ Voice*WordOrder + (Voice*WordOrder|Participant) + (Voice*WordOrder|Item))
lmer(logR5 ~ Voice*WordOrder + (0+Voice+WordOrder|Participant) + (0+Voice+WordOrder|Item))
lmer(logR6 ~ Voice*WordOrder + (0+Voice+WordOrder|Participant) + (Voice*WordOrder|Item))
lmer(logR7 ~ Voice*WordOrder + (0+Voice+WordOrder|Participant) + (Voice+WordOrder|Item))
References


https://doi.org/10.1098/rstb.2008.0310


https://doi.org/10.3758/BF03213769


3rd LFG Conference, Brisbane, Australia.


https://doi.org/10.3758/s13423-016-1040-7

https://doi.org/10.1007/BF00133368

Hahn, M., Degen, J., & Futrell, R. (2020, March). Modeling word and morpheme order in
natural language as an efficient tradeoff of memory and surprisal. 33rd Annual CUNY
Conference on Human Sentence Processing, Amherst, MA. https://psyarxiv.com/nu4qz

Second Meeting of the North American Chapter of the Association for Computational
Linguistics on Language Technologies*, 1–8.


*Behavioral and Brain Sciences, 33*(2–3), 61–83.
https://doi.org/10.1017/S0140525X0999152X

effects using pseudoword primes and a delayed lexical decision. *International Journal of

Tagalog*. Wilhelm Fink Verlag.

characteristics. In K. A. Adelaar & N. Himmelmann (Eds.), *The Austronesian languages of

Musgrave (Eds.), *Voice and grammatical relations in Austronesian languages* (Vol. 247, pp.

facilitates subsequent comprehension: Evidence from interleaved object naming in sentence


203


205


207


https://doi.org/10.1080/01690969738686835

http://arxiv.org/abs/cmp-lg/9409008


https://doi.org/10.1080/23273798.2015.1117117


Mishra, R., Singh, N., Pandey, A., & Huettig, F. (2012). *Spoken language-mediated anticipatory eye-movements are modulated by reading ability—Evidence from Indian low and high literates* [Application/pdf]. https://doi.org/10.16910/JEMR.5.1.3

https://doi.org/10.1111/j.1756-8765.2009.01072.x


https://doi.org/10.1016/j.jneuroling.2016.01.001


