PSYCHOLINGUISTIC AND SOCIOLINGUISTIC INFLUENCES ON THE PRODUCTION OF OPTIONAL PLURAL MORPHOLOGY AMONG BILINGUAL SPEAKERS OF YUCATEC MAYA AND SPANISH

A Thesis in Communication Sciences and Disorders by Lindsay Kay Butler-Trump

© 2016 Lindsay Kay Butler-Trump

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science

May 2016
The thesis of Lindsay Kay Butler-Trump was reviewed and approved* by the following:

Carol A. Miller  
Professor of Communication Sciences and Disorders  
Thesis Advisor, Interim Department Chair

Chaleece W. Sandberg  
Assistant Professor of Communication Sciences and Disorders

Nicole M. Etter  
Assistant Professor of Communication Sciences and Disorders

*Signatures are on file in the Graduate School.
Abstract

This is a study of the psycholinguistic and sociolinguistic factors that influence the use of optional plural morphology and number agreement among bilingual speakers of Yucatec Maya and Spanish. The conceptual factor of set size was determined to influence the production of optional plural morphology and agreement in this population. Speakers were more likely to use plural morphology on the noun and number agreement on the verb when a larger set size, e.g. seven versus two, was depicted. The opposite was found for the mention of a number word. Speakers were less likely to mention the number word in response to a larger set size. For this study, we recruited participants who range in age from 5 to 48 years old and who range in level of education from no formal education to some college education. Though age was not a predictor of the use of plural morphology or number agreement, speakers with higher levels of education produced significantly more plural morphology on nouns but not on verbs. A language-internal explanation is proposed to address this finding. I discuss the implications of these findings for populations served by Speech-Language Pathologists.
Table of Contents

List of Figures vi
List of Tables vii
Acknowledgments viii

Chapter 1
Introduction 1
1.1 Motivation ............................................. 1
1.2 Goals .................................................. 4
1.3 Previous experimental results and predictions of current study .... 5

Chapter 2
Background 7
2.1 Bilingualism in the Yucatan Peninsula ............................. 7
2.2 Yucatec Maya grammar and number marking ...................... 8

Chapter 3
Methods 12
3.1 Participants ............................................. 12
3.2 Procedure ............................................... 14
3.3 Materials ................................................ 15
3.4 Coding and exclusions .................................... 17

Chapter 4
Analyses 19
4.1 Set size .................................................. 20
4.2 Education and age ...................................... 21
4.3 Language use variables .................................. 22
4.4 Matching agreement .................................... 24
Chapter 5

Discussion

5.1 Conceptual effects of set size ........................................ 28
5.2 Processing of optional agreement .................................... 29
5.3 Effects of education on language use ............................... 30
5.4 Language-internal factors ............................................. 30
5.5 Less-commonly studied languages ................................. 31
5.6 Applications to clinical populations ............................... 32

Bibliography .................................................................. 35
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Location of the community of Popolá relative to the United States and Central America</td>
<td>13</td>
</tr>
<tr>
<td>3.2</td>
<td>Location of the community of Popolá relative to the Yucatan Peninsula</td>
<td>13</td>
</tr>
<tr>
<td>3.3</td>
<td>Age and level of education of participants</td>
<td>15</td>
</tr>
<tr>
<td>4.1</td>
<td>Proportion of plural marking on nouns by set size (error bars indicate bootstrapped 95% confidence intervals)</td>
<td>21</td>
</tr>
<tr>
<td>4.2</td>
<td>Proportion of plural marking on verbs by set size (error bars indicate bootstrapped 95% confidence intervals)</td>
<td>22</td>
</tr>
<tr>
<td>4.3</td>
<td>Proportion of plural marking on nouns by age (with linear smoothing in grey)</td>
<td>23</td>
</tr>
<tr>
<td>4.4</td>
<td>Proportion of plural marking on nouns by level of education (with linear smoothing in grey)</td>
<td>23</td>
</tr>
<tr>
<td>4.5</td>
<td>Proportion of matching plural agreement by condition (error bars indicate bootstrapped 95% confidence intervals)</td>
<td>25</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Cross-reference markers in Yucatec Maya ........................................... 9
3.1 Language use data including number of participants and approximate percentages .......................................................... 14
3.2 Example experimental stimuli by set size and animacy ............................ 16
3.3 Distribution of exclusions by set size ...................................................... 18
4.1 Hours per day using Spanish and use of number morphology ................. 24
4.2 Predictors and outcomes of mixed effects logit models .......................... 27
Acknowledgments

Funding for this research was provided by an Individual Grant from the Jacobs Research Funds, Whatcom Museum Foundation, Bellingham, Washington. Parts of this thesis were submitted for publication in Butler & Couoh Pool (under review). Many thanks are owed to Rosa Marfa Couoh Pool for linguistic and cultural consultation, recruiting participants, assisting in carrying out the experiment and coding the data. I would also like to thank Olga Esther Uc Cocom for assistance in running this experiment. This research would not have been possible without the help of the participants from the community of Popolá, whom I thank for their time and for inviting us into their homes to run the study. We thank audience members of the Center for Language Science Speaker Series at The Pennsylvania State University and at the Speech and Language Pathology and Audiology Brown Bag speaker series at Kent State University, where this research was presented. I owe many thanks to Carol Miller, Chaleece Sandberg, Nicole Etter and Carrie Jackson for helpful comments on this research. All errors are my own.
Chapter 1  
Introduction

1.1 Motivation

This is a study of the psycholinguistic and sociolinguistic factors that influence sentence processing in a less-commonly studied language, Yucatec Maya. A recent survey of how many different languages of the world have been studied in sentence production research suggests that a mere 0.6% of the world’s languages have been investigated with well-controlled psycholinguistic methods (Jaeger and Norcliffe, 2009). In Communication Sciences and Disorders as well, there is growing awareness of the fundamental role of cultural and linguistic diversity. ASHA (2013) outlines the ethical standards for professionals in serving the increasingly diverse population of the United States. Global engagement and international development are other areas of importance for the field, as outlined by Special Interest Group 17, Global Issues in Communication Sciences and Relateds Disorders.

Not only does this research contribute to general knowledge of cultural and linguistic diversity in the field, but the particular topic of the factors that influence the production of optional plural morphology can also be applied to clinical populations with the goal of increasing the effectiveness of intervention. There are many populations served by Speech-Language Pathologists and Audiologists who produce variable plural morphology and number agreement, for example, English language learners, children with expressive language delay, people with complex communication needs who use AAC, and adults with agrammatism. Typical adult speakers of English cannot provide information about how variable plural morphology and agreement are processed because the use of plural morphology and agreement is obligatory. Therefore, we cannot tell what factors will increase the use of plurals since they are always used. This is an area in which we can
look to different languages for an answer that English cannot give us. For speakers of Yucatec Maya, the use of plural morphology and number agreement is optional. Therefore, they can serve as a typical population on which to norm the production of variable plural morphology and agreement. The goal of this study is to investigate the psycholinguistic and sociolinguistic factors that affect the production of optional plural morphology and agreement among speakers of Yucatec Maya.

Across languages of the world, number is a common grammatical category (Corbett, 2000). For speakers of some languages, such as Spanish, when they utter a noun that refers to more than one entity, they must use plural morphology (e.g. *lo-s gato-s* ‘the-PL cat-PL’). Number agreement between the determiner and noun and between the subject and verb is similarly obligatory in Spanish. Other languages lack plural marking entirely, and the plurality of the noun is determined based on the context (and the grammar does not have number agreement). In Yucatec Maya, there is an optional plural morpheme (*-o’ob*), so if the plural morpheme is used (e.g. *le péek-o’ob* ‘the dog-PL’), it refers to more than one entity, but if no plural morpheme is used (e.g. *le péek* ‘the dog’), the noun can refer to either one or more than one entity (see below for more detail, and see Butler (2012) for a linguistic analysis of the nominal plural morpheme in Yucatec Maya). Similarly, number agreement between the noun and verb is optional (see Butler (2012)). Thus, Yucatec Maya-Spanish bilinguals present an interesting case since one language (Spanish) has obligatory plural marking and number agreement and the other has optional plural marking and number agreement (Yucatec Maya). In this study, we set the stage for the investigation of the processing of optional plural marking and number agreement among Yucatec Maya-Spanish bilinguals speaking in Yucatec Maya. The objective of this study is to examine the effects of psycholinguistic and sociolinguistic influences on the production of optional plural morphology among Yucatec Maya-Spanish bilinguals by carrying out a sentence production task in Yucatec Maya.

Apart from the theoretical importance of this research question, this line of research is relevant when applied to the field of communication sciences and disorders in a number of ways. There are many people with language disorders, differences or delays who may show variation in the use of plural morphology and number agreement. For example, a child who speaks African American English in the home speaks a language in which it is grammatical to say “Michael has four dog” without plural morphology when a number word is used (Bland-Stewart, 2005; Green, 2002; Labov, 1975; Seymour et al., 1998). Though there is little research on morphological development among multilingual children
who are acquiring English (but see Jia, 2003, on Mandarin-English-speaking children’s acquisition of plural morphology), there is evidence that the use of grammatical morphology is more variable among children who speak more than one language (Nicholls et al., 2011). Similarly, there is evidence that children who have moderate hearing impairment produce significantly less grammatical morphology, including nominal plural morphology (McGuckian and Henry, 2010). Even typically developing children under the age of 5 use grammatical morphology with some variation, and distinguishing typical language development from specific language impairment (SLI) is consequently challenging (Conti-Ramsden and Hesketh, 2010). Cross-linguistic research on the characteristics of SLI has shown variation across languages in the morphosyntactic characteristics of the disorder in Spanish (Bedore and Leonard, 2001) and Italian (Bortolini and Caselli, 1997). In addition, differences in processing grammatical morphology may depend on the richness of the morphological paradigm of the language among typical speakers (Bock et al., 2012; Lorimor et al., 2008) and among atypical speakers with SLI and aphasia (Bates et al., 1987).

Where does this leave us with respect to the investigation of the factors that affect the production of optional plural morphology among bilingual speakers of Yucatec Maya and Spanish? As discussed previously, typical adult speakers of English cannot provide much insight into the factors that affect variable plural morphology, since plural morphology is not highly variable (but there is some variability in the use of number agreement when a plural local noun attracts agreement on the verb, as in “The key to the cabinets are on the table” (Bock and Miller, 1991). A body of research in the processing of what is called the “agreement attraction effect” in English and other related languages grew out of this observation (Bock and Eberhard, 1993; Bock and Miller, 1991; Bock et al., 2012, 2004, 1999; Eberhard et al., 2005; Gillespie and Pearlmutter, 2011; Lorimor et al., 2008; Nicol et al., 1997; Vigliocco et al., 1996; Vigliocco and Nicol, 1998, inter alia). Since these studies are restricted to subject-verb number agreement with complex subjects, they cannot give us insight into variable plural marking in the nominal domain or in less complex sentence contexts. Since the use of plural morphology on nouns along with number agreement on verbs is optional in Yucatec Maya, it is a language that can uniquely answer the question of what factors influence the use of variable plural marking and number agreement. The objective of this study is to examine to what extent psycholinguistic factors (such as the conceptual effects of number information) and sociolinguistic factors (such as age, education and language use variables) affect the production of optional plural morphology and number agreement among bilingual speakers of Yucatec Maya and Spanish.
1.2 Goals

Little is known about bilingualism in the Yucatan Peninsula. Pfeiler and Zámiová (2006) classify bilingualism in the Yucatan Peninsula as individual level bilingualism with society-wide functional diglossia. Most speakers use Yucatec Maya in the home and in regional government affairs, and they use Spanish as the language of commerce and education. Since little is known about language use by Yucatec Maya-Spanish bilinguals, the first goal of this study is to investigate the role of education in the use of optional plural marking and number agreement among this population. We contrast the effects of education with the effects of age and the influence of language use variables, such as which language is spoken in the home and how many hours a day Spanish is spoken. However, since Spanish is the language of education, we expect education and proficiency in Spanish to be two highly correlated variables that we may not be able to disentangle. We also analyze participants’ use of Spanish based on their self-reports of how many hours per day they spend speaking Spanish. Due to the use of Spanish as the language of education, it may not be possible in this present study to disentangle the effects of Spanish use and education in Spanish.

Similarly, little is known about the processes that underlie the production of agreement that is not obligatory, since most psycholinguistic research on number agreement has been conducted with speakers of obligatory agreement languages (such as English and Spanish). In a psycholinguistic study of the production of optional plural marking with university-enrolled speakers of Yucatec Maya, Butler et al. (2014) found that conceptual number information influences the production of optional plural marking. In their study, a picture of seven entities was significantly more likely to lead to the use of plural morphology compared to a picture of two entities. Therefore, the second goal of this study is to replicate the findings of Butler et al. (2014) and then to extend their findings by analyzing these factors with mixed-effects logit regression analyses and by taking a more in-depth look at the effects of these variables on the production of number agreement between a predicate and an argument.
1.3 Previous experimental results and predictions of current study

Butler, Jaeger and Bohnemeyer (2014) conducted a picture description task with university-enrolled bilingual speakers of Yucatec Maya and Spanish with the goal of examining the conceptual effects of set size on sentence production in Yucatec Maya. They showed pictures of one, two and seven humans or animals depicting an intransitive actions (e.g. a girl sweeping or a chicken running). Among this highly educated population of speakers of Yucatec Maya, Butler et al. (2014) found that participants used significantly more plural morphology with larger set sizes (seven vs. two and two vs. one). In addition, speakers showed a preference for producing matching agreement, using plural agreement on the verb when the noun was also plural marked. This population of speakers is highly educated and thus highly proficient in Spanish (since Spanish is the language of wider education), so the influence of education (in Spanish) on the use of plural morphology and number agreement in Yucatec Maya is unclear. Due to Spanish being the language of education among this population of speakers, the relationship between education and proficiency in Spanish is one that is difficult to entangle. The goal of this study is to scratch the surface by examining the role of level of education on the use of optional plural morphology.

In this study, we employ the same materials as those used by Butler et al. (2014) to investigate the question of conceptual influences on the use of plural morphology and number agreement among a population of speakers of Yucatec Maya that are more diverse in terms of age and educational background. Speakers of Yucatec Maya in general are quite unique in that there are speakers with advanced degrees, and there are speakers with no formal education at all. In this study, the participants range in level of education from no formal education at all to an eleventh grade education. And, in our analyses, we pool the data from these participants with the university-enrolled participants from Butler et al. (2014). We intentionally recruited a pool of participants that range in age from 5 to 48 years old in order to rule out a correlation between age and education.

In addition to replicating the experiment conducted by Butler et al. (2014) with a larger pool of participants with more diverse demographics, we additionally address the question of the effects of high levels of education in Spanish on the production of optional plural marking and agreement. We also go beyond previous findings in that we use mixed effects models to analyze the effects of education and other demographic and language use
variables in addition to the effects of conceptual number information on the production of predicate-argument number agreement in Yucatec Maya.

We propose the following two competing predictions for our current study: 1) Conceptual number information will affect the production of plural marking and agreement regardless of education (since the language production mechanism is assumed to be universal and independent of sociolinguistic factors and language background), or 2) higher levels of education will lead to increased use of plural morphology since education is in Spanish, a language with obligatory plural marking and agreement.
Chapter 2  |  Background

Since Yucatec Maya is a less-commonly studied language, background information about bilingualism in the Yucatan Peninsula and about the grammar of Yucatec Maya and the system of optional plural marking and number agreement is presented below.

2.1 Bilingualism in the Yucatan Peninsula

Yucatec Maya is one of the most prominent languages in Mesoamerica. It is the language once spoken by the ancient Mayan civilization. Speakers of Yucatec Maya refer to themselves and their language as “maya,” whereas speakers of the other approximately 30 Mayan languages do not. The Instituto Nacional de Estadística, Geografía e Informática reports that there are 786,113 speakers of Yucatec Maya out of a total population of 2,057,753 over the age of 5 in the Yucatan Peninsula (the states of Campeche, Quintana Roo and Yucatan) (INEGI, 2010). Speakers of Yucatec Maya represent nearly 40% of the population of the peninsula. Among speakers of Yucatec Maya, there is widespread bilingualism in Spanish (or perhaps more accurately bilingualism (on an individual level) and functional diglossia (on a social/cultural level) (see Pfeiler and Zámiová, 2006, for an overview). Functional diglossia refers to the use of two languages, each restricted to particular formal situations. To provide an example of functional diglossia, it is common that speakers use Yucatec Maya in the home and for municipal matters in Maya-speaking communities. Speakers typically use Spanish in certain other contexts, such as in commerce, business and education. A bilingual education program initiated in 1955 by the Dirección General de Educación Indígena (DGEI) is considered largely responsible for the loss of indigenous languages in Mexico and for the decline in numbers of monolingual speakers of Yucatec Maya. Until around 1980, DGEI supported a focus on Hispanicization and assimilation for indigenous people in
Mexico. Starting in 1996, the Indigenous Intercultural Bilingual Education program (under DGEI) and the program for Educational Assistance to the Indigenous Population for Cultural Development (under the Consejo Nacional de Fomento Educativo (CONAFE, 2000)) represented a shifting focus toward teaching “conscious bilingualism” for the preservation of the Yucatec Maya language (Pfeiler and Zámiová, 2006). Changes are evident, though just beginning. In places such as La Universidad de Oriente, in Valladolid, Yucatán near where this study was conducted, students from indigenous Maya-speaking communities can major in Maya Linguistics and Culture and are placed in schools in indigenous communities to teach literacy and other subjects in the Maya language. Aside from special programs such as that, education is largely still conducted in Spanish. Therefore, it is difficult to disentangle the effects of education in Spanish and use of Spanish independent of level of education. In this study we examine the effects of level of education and also number of hours per day spent speaking Spanish, but we expect these variables to be highly correlated.

2.2 Yucatec Maya grammar and number marking

Yucatec Maya is a member of the Mayan language family. It is a head-marking, mildly polysynthetic language. In a head-marking language such as Yucatec Maya, grammatical relations are typically indicated by verbal affixes. Yucatec Maya has two sets of cross-reference markers that are realized as verbal prefixes and suffixes (and some discontinuous morphemes) (see Table 2.1). Set A references possessors and subjects of intransitive verbs in the imperfective aspects. Set B references subjects of intransitive verbs in the perfective aspects. In transitive sentences, Set A references the agent/subject, and Set B references the patient/object (see Bohnemeyer, 2002; Bricker, 1981). Due to this and the fact that the third person plural marker -o’ob is homophonous between Set A and Set B, there is some ambiguity that is inherent in the paradigm. For example, the sentence in (1) has multiple possible meanings depending on whether the morpheme -o’ob is interpreted as Set A (“They took it.”), Set B (“S/he took them.”) or both (“They took them.”). Moreover, the nominal plural marker -o’ob is also homophonous with (and possibly borrowed from) the verbal third person cross-reference marker -o’ob. Thus, the meaning of the phrase in (2) below also has multiple possible meanings depending on whether the morpheme -o’ob is interpreted as a Set A marker (“their dog”), a nominal plural marker (“his dogs”) or both (“their dogs”). See Butler (2012) for a syntactic analysis of the nominal plural morpheme
-o’ob as a determiner-level adjunct.¹

Table 2.1. Cross-reference markers in Yucatec Maya

<table>
<thead>
<tr>
<th>Person</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set A First</td>
<td>in-</td>
<td>k...o’on</td>
</tr>
<tr>
<td>Second</td>
<td>a(w)</td>
<td>a(w)...-e’ex</td>
</tr>
<tr>
<td>Third</td>
<td>u(y)</td>
<td>u(y)...(o’ob)</td>
</tr>
<tr>
<td>Set B First</td>
<td>-en</td>
<td>-o’on</td>
</tr>
<tr>
<td>Second</td>
<td>-ech</td>
<td>-e’ex</td>
</tr>
<tr>
<td>Third</td>
<td>-∅/-ij</td>
<td>-o’ob</td>
</tr>
</tbody>
</table>

(1) T-u bis-aj-o’ob
pfv-A3 caret-cmp-A3.pl/b3.pl
‘S/he took them.’/‘They took it.’/‘They took them.’ (Lucy, 1992, 53)

(2) u péek-o’ob
A3 dog-A3.pl/pl
‘his dog’/‘their dog’/‘their dogs’ (Lucy, 1992, 47)

Additionally, the use of the third person plural verbal cross-reference marker -o’ob is optional in the Set A paradigm and with nominal referents. This means that the sentence in (1) above without the plural morpheme -o’ob can still refer to a plurality of referents, as shown in (3) below.

(3) T-u bis-aj
pfv-A3 caret-cmp
‘S/he took it.’/‘S/he took them.’/‘They took it.’/‘They took them.’

Similarly, the phrase in (2) above can refer to a plurality of referents without the overt use of the plural morpheme -o’ob, as shown in (4) below.

(4) u péek
A3 dog
‘his dog’/‘his dogs’/‘their dog’/‘their dogs’

¹Abbreviations used include: A3 - Set A third person cross-reference marker, an - animate, b3-Set B third person cross-reference marker, cl-classifier, cmp-completive status marker, def - definite determiner, pfv-perfective aspect, pl-plural.
This system of verbal cross-reference and nominal plural marking is quite different from the better-known obligatory plural marking and agreement systems of many Indo-Eurasian languages, such as English and Spanish. Interestingly, our participants are bilingual speakers of Yucatec Maya and Spanish. Thus, they speak two languages with very different plural marking and agreement systems. In Spanish, when speaking about more than one object, the speaker must use a plural morpheme (e.g. *lo-s perro-s, ‘the-PL dog-PL’), but when speaking in Yucatec Maya, the speaker may choose to use or omit the plural morpheme (e.g. *le péek-o’ob, ‘the dog-PL, or le péek, ‘the dog’) and still refer to more than one dog. Additionally, when speaking in Spanish, the speaker must use plural agreement on the verb when the subject is also plural (e.g. *Lo-s perro-s ladr-an, ‘the-PL dog-PL bark-3PL’). In Yucatec Maya, however, the use of the plural morpheme is optional on the agent/subject noun as well as on the verb (see Butler, 2012).

Another relevant difference between Yucatec Maya and other more commonly studied Indo-European languages, such as English and Spanish, is that Yucatec Maya is a language in which speakers must use numeral classifiers when enumerating nominal referents. The example in (5) below shows that when talking about two dogs, a speaker of Yucatec Maya must include in the phrase, the noun, numeral and appropriate classifier. Typically, the noun ‘dog’ selects for the classifier for animate entities. The phrase in (5) without the classifier *túul is not grammatical (e.g. *ka’a péek).

(5) ka’a-túul péek
    two-CLAN dog
    ‘two dogs’

In Yucatec Maya, number agreement is not obligatory between the agent/subject and verb or between the noun and other nominal constituents, such as determiners and adjectives (Butler, 2012). This means that in the sentence in (6) below, a speaker could use the plural morpheme on the noun but not the verb, or on the verb but not the noun. Alternatively, the speaker could chose to mark neither the noun nor the verb with plural morphology, or the speaker could mark both the noun and verb with plural morphology. The speaker has similar plural marking and options in the phrase in (7) with a noun and adjective (Plural marking on determiners is not grammatical.).

(6) Táan u toj-ol(-o’ob) le péek(-o’ob)
    prog A3 yell-INC(-PL) def dog(-PL)
    ‘The dogs are barking’
The case of optional number agreement in Yucatec Maya is quite unlike Spanish, which has obligatory subject-verb and determiner-noun and adjective-noun agreement for number (among other grammatical features). This difference makes a very interesting case of Yucatec Maya-Spanish bilinguals since they speak two languages that differ in terms of number marking and agreement. We use field-based psycholinguistic methods (see Christianson and Ferreira, 2005)) to examine the processes that underlie the production of optional plural morphology and agreement. We test how conceptual number information influences the production of number agreement in addition to the use of plural morphology on nouns and verbs individually and on the mention of number words. We examine whether these effects hold regardless of education in Spanish along with other demographic and language use variables on the production of optional plural marking and agreement.
Chapter 3  |  Methods

3.1 Participants

We carried out the experiment in October of 2014 with 42 speakers of Yucatec Maya in the indigenous Maya-speaking community of Popolá, Yucatán, México, which lies 7 kilometers outside of the city of Valladolid in the center of the Yucatán Peninsula (see Figures 3.1 and 3.2). Twenty-one adults between the ages of 18 and 48 (20 females and 1 male) and 21 children between the ages of 5 and 12 (12 females and 9 males) participated in the experiment.\(^1\) The 42 participants ranged in age from 5 to 48 years old (\(M = 19.6, SD = 12.1\)), and they ranged in level of formal education from no formal education to an eleventh grade education (\(M = 5.1, SD = 2.5\) on a scale of 0 (no formal education) to 12 (final year of high school)). The experiment took no longer than 30 minutes, and the participants were compensated 50 Mexican pesos (about 4 U.S. dollars). In the case of participants under the age of 18, the parents were compensated 50 Mexican pesos and the children were given a number of small prizes (pens, stamps and notebooks), as was recommended by a native Yucatec Maya speaking consultant and a member of the community in which the experiment was conducted. Prior to taking part in the experiment, adult participants gave informed verbal consent in Yucatec Maya and child participants provided informed verbal assent in Yucatec Maya with parental consent as approved by the Institutional Review Board at the Pennsylvania State University.

We collected language background data from each participant requesting the following information: 1) all language(s) spoken (list), 2) language(s) spoken in the home (Maya,

\(^1\)We only had one adult male participant because it is common for females to work in the home and for males to work outside of the community, often going as far as the resort towns of the Peninsula, such as Cancun to find employment.
Figure 3.1. Location of the community of Popolá relative to the United States and Central America

Figure 3.2. Location of the community of Popolá relative to the Yucatan Peninsula

Spanish, both), 3) language(s) spoken at work/school (Maya, Spanish, both), 4) language(s) preferred (Maya, Spanish, both), 5) hours per day spent speaking Spanish (less than 1 hour, 1 to 3 hours, 3 to 6 hours, more than 6 hours). Table 3.1 summarizes the language use data from the 42 participants in this study. The most frequent answers were that participants
prefer both Maya and Spanish, speak mostly Maya in the home, speak mostly Spanish in
school or in the workplace, and spend an average of 1 to 3 hours per day speaking Spanish.

### Table 3.1. Language use data including number of participants and approximate percentages

<table>
<thead>
<tr>
<th>Hours speaking</th>
<th>Preferred language</th>
<th>Home language</th>
<th>School/work language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 1</td>
<td>Maya 17 (41%)</td>
<td>30 (72%)</td>
<td>5 (12%)</td>
</tr>
<tr>
<td>1 to 3 hours</td>
<td>Spanish 1 (2%)</td>
<td>1 (2%)</td>
<td>27 (64%)</td>
</tr>
<tr>
<td>3 to 6 hours</td>
<td>Both 24 (57%)</td>
<td>11 (26%)</td>
<td>10 (24%)</td>
</tr>
<tr>
<td>more than 6 hours</td>
<td>3 (7%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In addition to the participants recruited for this study, in our analyses, we pool the
experimental responses from the Yucatec Maya-Spanish bilingual students enrolled at La
Universidad de Oriente in Valladolid, Yucatán, México who participated in the Butler et al.
(2014) study. The university-enrolled participants were between the ages of 19 and 26 (14
females and 13 males). This led to a total of 69 participants. Figure 3.3 shows the age and
level of education of all 69 participants (with shapes indicating participant group: college
student at La Universidad de Oriente, child in Popolá and adult in Popolá).

### 3.2 Procedure

Two Yucatec Maya-Spanish bilingual consultants and members of the community of Popolá
conducted all background surveys, gave experimental instructions and carried out the
experiment in Yucatec Maya. After giving informed consent, participants were given the
language background survey. The experiment was delivered on an iPad with the slideshow
app Slideshark (Brainshark, Inc., 2014). Participants were instructed to look at the picture
on the iPad and say one sentence to describe what was there and what it was doing. Then,
the participants were given four practice items followed by an opportunity to ask any
questions. If a participant did not give an appropriate response, the consultant would
give an example response and let the participant practice again. The consultants and the

---

2The differences between the methods used by Butler et al. (2014) and the current study are: 1) Butler
et al. (2014) delivered the experiment on a laptop with experimental software in a university computer lab,
and the current study was conducted on an iPad with a slideshow app in the homes of the participants and 2)
Four of the experimental stimuli from Butler et al. (2014) were removed and replaced because these stimuli
tended to elicit transitive rather than intransitive responses and 3) Butler et al. (2014) gave instructions to
participants in Spanish, and we gave instructions to participants in Yucatec Maya.
participant sat at a small table (provided by the participant in their home). The consultants sat across the table from the participant so that they could not see the picture items on the iPad. The iPad was propped up and angled toward the participant for ease of viewing. Both consultants independently live-coded the experimental responses for: 1) mention of a numeral, 2) use of plural morphology on the noun, 3) use of plural morphology on the verb. After coding a response, one consultant controlled the swiping gesture on the iPad that advanced the experiment to the next trial. The experimental sessions were audio recorded.

### 3.3 Materials

The materials used in this experiment are based on the picture stimuli developed by Butler et al. (2014). The stimuli include twenty-four black and white sketch-like pictures that were chosen to elicit a sentence with a noun and intransitive verb (e.g. *The frogs are jumping*) (see Table 3.2). Twelve of the stimulus pictures depicted humans and twelve depicted animals. Four of the stimuli from Butler et al. (2014) were replaced (two of humans and two of animals) with four alternate pictures because these items had elicited many transitive (subject-verb-object) responses in their experimental responses.

The twenty-four items were pseudo-randomized in a Latin Squares design so as to present a picture of a human followed by a picture of an animal and to ensure that no...
participant saw the same item more than once. There were twenty-four fillers for adult participants and twelve fillers for child participants (fewer fillers for children in order to ensure they could complete the experiment). The fillers depicted humans and animals involved in transitive actions, such as a girl playing a guitar. The fillers varied the number of referents that the subject and object nouns depicted (e.g. three children boarding a bus or a mouse eating three pieces of cheese). The potential responses are shown below based on the sentence “The dog is barking.” The sentence in 8) has no plural morphology. The sentence in (9) has plural morphology on the noun only. The sentence in (10) has plural morphology on the verb only. The sentence in (11) has plural morphology on the noun and verb. In addition to these responses, the participant could also decide to include the numeral, saying for example, “One dog is barking”, “Two dogs are barking” or “Seven dogs are barking” in combination with the plural marking possibilities in (8) through (11) below.

(8) Le péek táan u toj-ol
    def dog   prog a3 yell-inc
    ‘The dog is barking’ / ‘The dogs are barking’
3.4 Coding and exclusions

Responses for the 24 experimental items from the 27 university-enrolled participants from Butler et al. (2014) and the 42 participants recruited for this study led to a total of 1,656 responses from 69 participants. Responses were excluded if they were considered “uncodable” (i.e. the recording was cut off, the participant gave no response, the response was unintelligible, or the coder was unsure of a relevant aspect of the response). Responses that were missing a constituent (e.g. no noun or no verb mentioned) were excluded. Responses that were verb-initial (only 3.3% of the total responses) were excluded due to existing evidence that verbs in verb-initial sentences are significantly less likely to be marked with plural morphology than verb-final sentences (Butler, 2011). Finally, responses in which a transitive verb with an object noun was mentioned (8.1% of the total responses (e.g. “The pig is eating slop” rather than “The pig is eating”)) were excluded in case the presence of an overt object noun phrase affected the likelihood of optional plural marking on the verb.3 There was no significant difference in exclusion rates across conditions (uncodable data: $\chi^2(2) = 2.6, p = 0.27$, missing constituents: $\chi^2(2) = 0.97, p = 0.62$, verb-initial responses: $\chi^2(2) = 0.15, p = 0.93$, object mentioned: $\chi^2(2) = 4.2, p = 0.12$). The distribution of excluded data by reason for exclusion is shown in Table 3.3 below.

---

3The majority of exclusions due to object-mention were from the university-enrolled participants first reported in Butler et al. (2014) since the four items that frequently led to object mention were removed and replaced in the current study.
Table 3.3. Distribution of exclusions by set size

<table>
<thead>
<tr>
<th>Set size</th>
<th>Uncodable</th>
<th>Missing constituent</th>
<th>Verb-initial</th>
<th>Object mentioned</th>
<th>Remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>20 (3.6%)</td>
<td>12 (2.2%)</td>
<td>17 (3.1%)</td>
<td>51 (9.2%)</td>
<td>452 (81.9%)</td>
</tr>
<tr>
<td>Two</td>
<td>28 (5.1%)</td>
<td>16 (2.9%)</td>
<td>19 (3.4%)</td>
<td>45 (8.2%)</td>
<td>444 (80.4%)</td>
</tr>
<tr>
<td>Seven</td>
<td>32 (5.6%)</td>
<td>17 (3.1%)</td>
<td>17 (3.1%)</td>
<td>33 (6%)</td>
<td>453 (82.2%)</td>
</tr>
<tr>
<td>Total</td>
<td>80 (4.8%)</td>
<td>45 (2.7%)</td>
<td>53 (3.2%)</td>
<td>129 (7.8%)</td>
<td>1349 (81.5%)</td>
</tr>
</tbody>
</table>
Chapter 4 | Analyses

We conducted mixed effects logit regression analyses using R (R Development Core Team, 2008) and the lme4 package (Bates et al., 2015a,b) in order to determine the relationship between predictor and outcome variables while factoring in to the same model the random effects of individual subjects and items (Breslow and Clayton, 1993; Jaeger, 2008). This is an appropriate statistical method considering the design of the experiment and the potential for individual variation in the use of optional linguistic elements. All mixed effect logit models include random intercepts by subjects and items but not random slopes, since including random slopes led to over-parameterization of the models. In all analyses reported collinearity was not an issue (all predictors were centered with fixed effects rs < .6).

The outcome variables plural noun, plural verb, numeral mention and plural agreement match (covariation of plural form on the subject and verb) were analyzed in four separate models (see Table 4.2 for a summary of all models). The main predictors in this analysis are education, age and set size. Numeral mention was also included as a predictor variable in the models with the outcomes plural noun, plural verb and plural agreement match. The continuous variables education and age were centered around the mean. Categorical variables were either Helmert coded or treatment coded (user-defined). Helmert coding was used to compare the levels of a variable with the mean of subsequent levels. The categorical variable set size was Helmert coded for the analysis of plural marking and treatment coded for the analysis of the mention of numerals. For the analysis of plural morphology use on nouns and verbs, the variable set size was Helmert coded to examine effects of one versus more than one and two versus seven. For the analyses of the mention of numerals, the variable set size was treatment coded to contrast two versus one and two versus seven. The motivation for treatment coding (versus Helmert coding) for numeral mention is that in
Yucatec Maya, the numeral one (e.g. the numeral and classifier combination, *jun-túul*, for one animate entity) is homophonous with the indefinite article. Thus, it is not possible to determine whether a numeral or an indefinite determiner was intended in the one condition. For this reason, we chose the set size of two as the baseline for the comparison. Moreover, there is existing evidence that the processing of numerals is facilitated for smaller non-singleton set sizes (i.e. two) versus larger ones (four or six) in studies of English speaking children acquiring nominal plural marking (Barner et al., 2012; Lanter and Basche, 2014; Zapf and Smith, 2008).

### 4.1 Set size

We examined the effects of conceptual information via the manipulation of set size, comparing sets of one, two and seven entities in this experiment. We examined the effects of set size on the use of plural morphology on nouns and verbs as well as the mention of number words (e.g. a response of “Two frogs are jumping” vs. “The frogs are jumping”). On nouns, plural morphology was used at an overall rate of 37.2% in the two condition and 56.3% in the seven condition. On verbs, plural morphology was used at an overall rate of 59.9% in the two condition and 71.7% in the seven condition. Set size was a significant predictor of the use of plural morphology on nouns and verbs (*p* < .001). For nouns and verbs, participants used significantly more plural morphology with set sizes of two and seven compared to one, and they used more plural marking with a set size of seven compared to two (see Table 4.2 for details of the mixed effects models). Figures 4.1 and 4.2 show the proportion of plural morphology used on nouns and verbs with increasing set size.

Numerals were mentioned at an overall rate of 40.8% in the one condition, 52.2% in the two condition and 39.7% in the seven condition. Set size had a significant effect on the mention of a numeral, but the effect did not go in the same direction as the effect of set size on the use of plural morphology on nouns and verbs. Participants were significantly less likely to mention the numeral when shown a picture with a set size of one compared to a set size of two, and they were significantly less likely to mention the numeral with a set size of seven compared to a set size of two (*p* < .001) (see Table 4.2 for the details of the mixed effects models). Participants were most likely to mention the numeral when describing smaller non-singleton sets.

In addition, set size and numeral mention interacted in the production of plural morphology on nouns and verb and in the production of plural agreement. When no numeral
was mentioned (vs. when a numeral was mentioned), the rate of plural marking on nouns was significantly greater in the seven condition but not in the two condition ($p < .01$) (see Figure 4.1). The same interaction effect held for the production of matching plural agreement. When no numeral was mentioned (vs. when a numeral was mentioned), the rate of matching plural agreement was significantly greater in the seven condition but not in the two condition ($p < .01$) (see Figure 4.5) below. For the production of plural morphology on verbs, however, the interaction effect went in a different direction. When no numeral was mentioned (vs. when a numeral was mentioned), the rate of plural morphology used on verbs was significantly lower in the two condition while there was no difference in rate of plural use on verbs depending on numeral mention in the seven condition ($p < .001$) (see Figure 4.2).

**Figure 4.1.** Proportion of plural marking on nouns by set size (error bars indicate bootstrapped 95% confidence intervals)

### 4.2 Education and age

In addition to set size, we considered the effects of level of education on the use of optional plural morphology and numeral mention. We also analyzed the effects of age in order to rule out a potentially confounding effect on level of education. Age did not significantly
predict the use of plural morphology on the noun or the verb (see Figure 4.3), nor did it significantly affect the production of matching plural agreement. Similarly, age did not significantly predict the mention of numerals (see Table 4.2 for full details of the mixed effects models). On the other hand, level of education significantly predicted the use of plural morphology on nouns \((p < .05)\) (see Figure 4.4), but it did not predict the use of plural morphology on verbs or the mention of a numeral. Level of education also significantly predicted the production of matching plural agreement \((p < .05)\). Higher levels of education resulted in significantly higher rates of use of the plural on nouns as well as higher rates of matching plural agreement.

### 4.3 Language use variables

The following language background information was collected from participants in this study: 1) language spoken (list), 2) language(s) spoken in the home (Maya, Spanish, both), 3) language(s) spoken at work/school (Maya, Spanish, both), 4) language(s) preferred (Maya, Spanish, both), 5) hours per day spent speaking Spanish (less than 1 hour, 1 to 3 hours, 3 to 6 hours, more than 6 hours) (refer back to Table 3.1 for a summary). Since these data were not designed as predictors of the outcome variables analyzed in the mixed
effects logit regression models, they are not appropriate for such an analysis. Additionally, due to the low cell counts and highly probable correlations among these variables, they are not appropriate for a mixed-effects logit analysis. For example, low cell counts occurred because only one speaker reported that he or she speaks only Spanish at home. Due to zero cell counts for home language and preferred language (only one participant answering Spanish for both categories), we focus on participants’ estimates of how many hours per day they spend speaking Spanish. In a Chi-squared test, the distribution of plural nouns produced was significantly higher as the number of hours per day speaking Spanish
increased ($\chi^2(3) = 10.63, p < .05$) (and see Table 4.1). The same effect did not hold for the production of plural morphology on verbs ($\chi^2(3) = 5.53, p = .14$), though the effect went in the same direction (see percentages of plural verb use in Table 4.1). Similarly, there was a significant effect of hours per day speaking Spanish on the production of matching plural agreement (covarying plural morphology used on both the noun and verb) ($\chi^2(3) = 10.34, p < .05$). This finding aligns with the result of the effects of education of the use of plural morphology. Higher levels of education led to significantly higher rates of use of plural morphology on the noun and on matching plural agreement. Likewise, an increasing number of hours per day a speaker of Yucatec Maya uses Spanish led to higher rates of use of plural morphology on nouns and matching plural agreement. This result suggests an effect in the same direction of education in Spanish and amount of time spent speaking Spanish. These variables may not be perfectly correlated, however, since there may be individuals who speak in Spanish many hours per day but don’t have a high level of education or individuals who are highly educated but spent more of their time speaking Maya than Spanish. A Pearson product-moment correlation coefficient was computed to assess the relationship between level of education and hours per day speaking Spanish in this data set. There was a positive correlation between the two variables ($r = 0.48, n = 945, p < .001$) indicating an increase in amount of time spent speaking Spanish with an increase in level of education.

<table>
<thead>
<tr>
<th>Hours/day speaking Spanish</th>
<th>Less than 1 hour</th>
<th>1 to 3 hours</th>
<th>3 to 6 hours</th>
<th>More than 6 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plural morphology on noun</td>
<td>22%</td>
<td>29%</td>
<td>37%</td>
<td>38%</td>
</tr>
<tr>
<td>Plural morphology on verb</td>
<td>40%</td>
<td>41%</td>
<td>49%</td>
<td>54%</td>
</tr>
<tr>
<td>Agreement match</td>
<td>21%</td>
<td>27%</td>
<td>35%</td>
<td>35%</td>
</tr>
</tbody>
</table>

### 4.4 Matching agreement

Overall, speakers produced sentences with no plural marking on the noun or verb (N-∅ V-∅) at a rate of 54%. Speakers produced plural morphology on both the noun and verb (N-pl V-pl) at a rate of 29.8%. Speakers produced agreement mismatches with plural marked only on the noun (N-pl V-∅) at a rate of 1.7%, and they produced mismatches with plural morphology used only on the verb (N-∅ V-pl) at a rate of 14.5%.
For the conditions in which the use of plural morphology would be semantically appropriate (two and seven, but not one), the production of plural morphology on both the noun and verb occurred at a rate of 35.8% in the two condition and 53.0% in the seven condition. Participants produced plural agreement mismatches (using plural morphology on the noun but not the verb or vice versa) at an overall rate of 25.5% in the two condition and 22.1% in the seven condition (see Figure 4.5 below, which shows the proportion of matching plural agreement produced). The Spearman’s rank order correlation coefficient (i.e. Spearman’s rho), a nonparametric procedure, was performed to address speakers’ preferences to produce matching agreement (plural morphology used on both the noun and verb or on neither). The Spearman’s rho revealed a significant relationship between the use of plural morphology on nouns and verbs ($\rho_{[1287]} = .69, p < .001$). Replicating Butler et al. (2014), we found that speakers significantly preferred to produce matching agreement overall.

![Figure 4.5. Proportion of matching plural agreement by condition (error bars indicate bootstrapped 95% confidence intervals)](image)

Going beyond Butler et al. (2014), we conducted mixed effects logit regression analyses with the variables set size and numeral mention predicting the production of matching plural agreement (covariation of plural form used on the subject and verb). Participants produced
significantly more matching plural agreement in the two and seven conditions compared to the one condition \((p < .001)\), as would be expected. In addition, participants produced significantly more matching plural agreement in the seven condition compared to the two condition \((p < .001)\), a significant effect of set size on the production of matching agreement. There was no significant effect of numeral mention of the production of matching plural agreement, but there was a significant interaction with set size. In the seven condition, but not the two condition, when no numeral was mentioned, participants produced significantly more matching plural agreement \((p < .01)\) (see Figure 4.5) and Table 4.2 for details of the mixed-effects models.
## Table 4.2. Predictors and outcomes of mixed effects logit models

<table>
<thead>
<tr>
<th>Outcome: Plural noun</th>
<th>β</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predictors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.86</td>
<td>.26</td>
<td>-7.15</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Set size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One vs. Two, Seven</td>
<td>-3.38</td>
<td>.31</td>
<td>-11.04</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Two vs. Seven</td>
<td>-.54</td>
<td>.09</td>
<td>-6.10</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Numeral Used</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeral used</td>
<td>.04</td>
<td>.03</td>
<td>1.15</td>
<td>= .25</td>
</tr>
<tr>
<td>One vs. Two, Seven:Numeral Used</td>
<td>-.18</td>
<td>.04</td>
<td>-4.24</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Two vs. Seven:Numeral Used</td>
<td>-.12</td>
<td>.05</td>
<td>-2.51</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Age, Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.31</td>
<td>.04</td>
<td>8.08</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Age</td>
<td>.00</td>
<td>.00</td>
<td>-0.6</td>
<td>= .95</td>
</tr>
<tr>
<td>Education</td>
<td>.01</td>
<td>.01</td>
<td>2.6</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Age:Education</td>
<td>-.00</td>
<td>.00</td>
<td>-0.6</td>
<td>= .96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome: Plural verb</th>
<th>β</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predictors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.45</td>
<td>.02</td>
<td>18.11</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Set size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One vs. Two, Seven</td>
<td>-.64</td>
<td>.02</td>
<td>-31.99</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Two vs. Seven</td>
<td>-.13</td>
<td>.02</td>
<td>-5.52</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Numeral Used</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeral Used</td>
<td>-.11</td>
<td>.03</td>
<td>-3.78</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>One vs. Two, Seven:Numeral Used</td>
<td>-.06</td>
<td>.04</td>
<td>1.54</td>
<td>= .12</td>
</tr>
<tr>
<td>Two vs. Seven:Numeral Used</td>
<td>-.19</td>
<td>.05</td>
<td>-4.18</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Age, Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.42</td>
<td>.00</td>
<td>10.03</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Age</td>
<td>.00</td>
<td>.00</td>
<td>.38</td>
<td>= .71</td>
</tr>
<tr>
<td>Education</td>
<td>.01</td>
<td>.01</td>
<td>1.9</td>
<td>= .06</td>
</tr>
<tr>
<td>Age:Education</td>
<td>-.00</td>
<td>.00</td>
<td>-0.3</td>
<td>= .98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome: Numeral used</th>
<th>β</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predictors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.4</td>
<td>.05</td>
<td>8.03</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Set size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two vs. One</td>
<td>-.07</td>
<td>.02</td>
<td>-4.08</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Two vs. Seven</td>
<td>.12</td>
<td>.02</td>
<td>6.12</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Age, Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.41</td>
<td>.05</td>
<td>7.94</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Age</td>
<td>.01</td>
<td>.01</td>
<td>.85</td>
<td>= .4</td>
</tr>
<tr>
<td>Education</td>
<td>.02</td>
<td>.01</td>
<td>-1.34</td>
<td>= .18</td>
</tr>
<tr>
<td>Age:Education</td>
<td>.00</td>
<td>.00</td>
<td>.48</td>
<td>= .63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome: Plural match</th>
<th>β</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predictors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.31</td>
<td>.02</td>
<td>12.43</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Set size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One vs. Two, Seven</td>
<td>-.41</td>
<td>.02</td>
<td>-19.1</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Two vs. Seven</td>
<td>-.15</td>
<td>.02</td>
<td>-6.08</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Numeral Used</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeral Used</td>
<td>.04</td>
<td>.03</td>
<td>1.23</td>
<td>= .22</td>
</tr>
<tr>
<td>One vs. Two, Seven:Numeral Used</td>
<td>-.18</td>
<td>.04</td>
<td>-4.24</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Two vs. Seven:Numeral Used</td>
<td>-.12</td>
<td>.05</td>
<td>-2.51</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Age, Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.29</td>
<td>.04</td>
<td>7.86</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Age</td>
<td>.00</td>
<td>.00</td>
<td>.01</td>
<td>= 1</td>
</tr>
<tr>
<td>Education</td>
<td>.02</td>
<td>.01</td>
<td>2.52</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Age:Education</td>
<td>.00</td>
<td>.00</td>
<td>.01</td>
<td>= 1</td>
</tr>
</tbody>
</table>
Chapter 5  |  Discussion

We investigated the effects of conceptual number information on the use of plural morphology, numeral mention and number agreement among bilingual speakers of Yucatec Maya and Spanish with varying levels of education in Spanish. In a field-based psycholinguistic sentence production task, we manipulated set size as a predictor of the use of optional plural morphology. We also examined the effects of education and other demographic and language use variables. The goal of this study was two-fold: 1) to elucidate the processes that underlie sentence production in a less-commonly studied language (with optional plural marking and number agreement) and 2) to investigate the effects of education in Spanish on the use of optional plural marking and number agreement among bilingual speakers of Yucatec Maya.

5.1 Conceptual effects of set size

Conceptual number information has a significant effect on the production of optional plural marking on nouns and verbs, the mention of numerals, and the production of matching agreement among bilingual speakers of Yucatec Maya and Spanish. Interestingly, how conceptual information affects the production of different linguistic elements diverges. Larger set sizes (seven vs. two) lead to higher rates of plural morphology use but lower rates of numeral mention. This finding aligns with research on the acquisition of plural marking by children who speak obligatory plural marking languages. Zapf and Smith (2008) and Barner et al. (2012) found that English speaking children between the ages of 2 and 4 who are in the process of acquiring plural morphology produced more plural morphology with larger set sizes of (e.g. four vs. two). They produced the most number words for set sizes of two compared to one, three, four, and higher numbers. Where English-acquiring
children diverge from speakers of optional-plural languages is that the children in the Barner et al. (2012) study were more likely to use plural marking when they had also mentioned a numeral. As discussed by Butler et al. (2014) our results indicated that for speakers of optional plural marking languages, plurality cues trade-off since speakers were less likely to use a plural morpheme on the noun when they had mentioned a numeral. These findings suggest a universal effect of conceptual number information based on set size that leads to divergent effects on the production of plural morphology and number words. Furthermore, this finding that conceptual number information affects the production of plural morphology and agreement with adult speakers is a result that could not have been determined based solely on adult speakers of a more commonly studied language such as English or Spanish, since the production of plural morphology and number agreement are obligatory.

5.2 Processing of optional agreement

The results of this study also set the stage for future investigations of the real-time computation of agreement with speakers of optional agreement languages. One phenomenon central to the theory of agreement processing is the “agreement attraction effect.” In the first psycholinguistic investigation of this effect, Bock and Miller (1991) found that speakers of English are more likely to produce an agreement mismatch with a singular matrix noun and a local plural “distractor” noun, such as “The key to the cabinets are on the table” (and see discussion of this phenomenon by linguists (e.g. den Dikken, 2001; Francis, 1986; Jespersen, 1924; Kimball and Aissen, 1971; Quirk et al., 1985, *inter alia*)). The agreement attraction effect has been reliably demonstrated in psycholinguistic experiments focusing on the processing pressures involved language production (Bock and Cutting, 1992; Bock and Eberhard, 1993; Bock and Miller, 1991; Gillespie and Pearlmutter, 2011; Staub, 2009, 2010; Vigliocco et al., 1996; Vigliocco and Nicol, 1998). Similarly, subject-verb agreement mis-matches are known to cause comprehension difficulty as well (Kreiner et al., 2013; Osterhout and Mobley, 1995; Pearlmutter et al., 1999). And, the agreement attraction affect has been demonstrated in sentence comprehension; plural-marked local nouns can interfere with the computation of verb agreement during comprehension (Kaan, 2002; Nicol et al., 1997; Pearlmutter, 2000; Pearlmutter et al., 1999; Thornton and MacDonald, 2003; Wagers et al., 2009). These findings have led researchers to propose that the mechanisms that underlie production and comprehension may be shared. There is, however, a fundamental confound involved in studying the agreement attraction effect with speakers of languages
for which agreement is obligatory – agreement mismatches in languages with obligatory agreement are erroneous productions, even though they may be produced by speakers. The above cited evidence that agreement attraction disrupts comprehension supports this notion, since speakers detect that an error has been produced. Speakers of languages in which there is optional agreement, such as Yucatec Maya, can shed new light on the question of whether the processes fueling the agreement attraction effect are specific to the production of errors and/or the detection of errors in comprehension. We hope future studies with speakers of Yucatec Maya will address this fundamental question.

5.3 Effects of education on language use

Since there is currently little known about bilingual speakers of Yucatec Maya and Spanish, especially through the lens of well-controlled psycholinguistic studies, we asked whether education and other demographic and language use variables influenced the use of optional plural morphology, numeral mention and agreement matching. Since Spanish is the language of education in the Yucatan Peninsula, even though 40% of the population speaks Yucatec Maya, disentangling the effects of education and use of Spanish was not possible. Education in Spanish and use of Spanish (regardless of level of education) were highly correlated. In our analysis of level of education, we found a significant effect of education on the production of optional plural morphology, but only for nouns. Speakers with higher levels of education produced more plural morphology, but only on nouns. Similarly, we found an effect of use of Spanish in that the more hours per day a participant used Spanish, the more plural morphology the individual produced, but only on nouns. The intriguing question that arises based on these results is why would these factors not significantly influence the use of plural morphology on verbs? In the next section, we offer a language-internal explanation for this otherwise puzzling finding.

5.4 Language-internal factors

Why is it the case the level of education and hours per day speaking Spanish predict the use of plural morphology on nouns but not verbs? An explanation lies in a deeper understanding of the grammar of Yucatec Maya. Yucatec Maya is a head-marking type language (Bohnemeyer, 2002, 2009; Bricker, 1981; Norcliffe, 2009) in which grammatical relations are marked on the heads (verbs) rather than the dependents (nouns) (Nichols,
There is evidence that syntactic relations between the verbal head marked with cross-reference morphology and a cross-referenced noun phrase may be different than that of agreement in dependent-marking languages (Austin and Bresnan, 1996; Bresnan and Mchombo, 1987; Jelinek, 1984; Van Valin, 1977, 1985), but for some head-marking languages this distinction is less clear. For many head-marking type languages, it seems to be the case that cross-reference markers saturate the syntactic argument positions in case no co-indexed noun phrase is present (e.g. Bresnan and Mchombo, 1987). And, there is evidence that this is indeed the case for Yucatec Maya (Bohnemeyer et al., 2015).

What this fact about the grammar of Yucatec Maya would mean for language processing is that grammatical relations may occasionally be processed when the speaker produces or comprehends the first noun phrase (in a subject initial utterance), but they are more reliably processed at the verb. In other words, there are sometimes cues to the grammatical role of the sentence-initial noun within the noun phrase, but they are optional. Grammatical markers indicating person and number are more reliably marked on the verb. What this implies is that speakers put more weight on verbal morphology than nominal morphology, and therefore, the use of optional nominal plural morphology on the noun is more susceptible to the influence of education in Spanish and use of Spanish among bilingual speakers. There is emerging evidence that head-marking languages, such as Yucatec Maya, are processed differently from dependent-marking languages (Norcliffe and Jaeger, 2014; Norcliffe et al., 2015; Norcliffe, 2009) and more generally that the structure of a language may affect early language production processes (e.g. Brown-Schmidt and Konopka, 2008; Sauppe et al., 2013).

5.5 Less-commonly studied languages

This study contributes to the empirical base of an understudied language, Yucatec Maya. And, this study contributes to a phenomenon that is understudied among Mayan languages, number marking and agreement (see England, 2011). Importantly, this study may have implications for the maintenance and revitalization of endangered languages. For an indigenous language such as Yucatec Maya, speakers frequently see borrowings and superstrate influences enter into their language from a language of wider communication, such as Spanish. As a result of this study, we found that higher levels of education in Spanish (though highly correlated with language use regardless of level of education) predicted the increased use of optional plural morphology on nouns. This result suggests that education...
in Yucatec Maya may be an important factor in maintaining and revitalizing the language, though more research on this topic would be required to answer this question. Although national educational objectives exist in Mexico for bilingual education that incorporates indigenous languages as shared language of instruction (e.g. the aforementioned Consejo Nacional de Fomento Educativo), among our community of speakers of Yucatec Maya in Popolá, and likely for other indigenous communities around the Yucatan Peninsula and other parts of Mexico, this educational objective has not been implemented in practice.

5.6 Applications to clinical populations

As discussed in the introduction, there are many clinical populations who may be producing variable plural morphology and number agreement to whom this research may apply. Some clinical populations that may be producing variable plural marking include adults with agrammatic aphasia, children with language delay or impairment, English language learners, and people with complex communication needs who use AAC. For example, a child client who speaks African American English speaks a language in which the use of plural morphology is optional when a numeral is used. If this child has a goal of increasing the use of plural morphology in Standard American English, the results of this research are particularly relevant. First, in order to increase the use of plural morphology, the clinician would choose stimuli such as pictures that showed larger set sizes. A clinician would want to show a picture of seven dogs rather than two. Not only would we expect the larger set size to elicit more plural morphology, but also it should suppress the use of a numeral. In African American English, the use of a number word renders the use of plural morphology optional (Bland-Stewart, 2005; Green, 2002; Labov, 1975; Seymour et al., 1998), so the clinician would want to avoid the child using number words in this case. The same implications would hold for a client who is an English language learner (particularly if the client’s first language lacks plural marking) or a child who has an expressive language delay or language impairment who has a goal of increasing the use of plural morphology. This research is particularly relevant to clients with complex communication needs who use picture-based AAC systems since grammatical morphemes are typically overlooked in such systems (Binger and Light, 2008). The application based on the findings of this project is that the visual representation of plural should involve a larger set size, rather than a depiction of just two items. This research may also apply to adult clients with agrammatism who have a goal of increasing the use of plural morphology and number agreement, but this
line of research would have to consider whether the underlying mechanism of the deficit is due to a faulty conceptual representation or due to a faulty syntactic representation and/or processing deficit.

Where does that leave us with respect to how the results of this study could lead to better intervention outcomes? Strong evidence exists that complexity enhances learning in typical communication as well as disordered communication. Speakers who are exposed to more acoustic-phonetic variability learn better. For example, adults exposed to talkers from three different dialect varieties adapted to unfamiliar talkers better than participants who were exposed to a talkers from a single dialect variety (Clopper and Pisoni, 2004). Similarly, infants who are exposed to more phonetic variability perform better on word learning tasks (Rost and McMurray, 2010). Typical language acquisition in general has been shown to arise as a result of exposure to complex linguistic input rather than simplified speech and language patterns (Newport et al., 1973, 1977). In the field of communication disorders, the Complexity Account of Treatment Efficacy (CATE) (Thompson, 2007; Thompson et al., 2003) makes the explicit proposal that training complex structures, rather than simpler related structures, promotes generalization in the treatment of language disorders. The complexity account has been shown to account for better treatment outcomes in the domain of phonological learning, word learning and syntactic learning. Gierut (2007) discusses numerous examples of improved treatment outcomes for phonological disorders as a result of treatment with complex phonological targets, including phonetic and phonemic inventories, distributional properties of sounds, syllable structure and phonological processes.

Similarly, Kiran (2007) reviews evidence that treatments that target semantically complex words result in better intervention outcomes. In particular, training atypical vs. typical items within a category, improves naming outcomes for individuals with aphasia. The stimuli used in this study align with this view of complexity in that items that depicted a small set size are a subset of the items depicting a large set size. That is to say that a set of two is a subset of a set of seven. A set size of two items represents the minimum set size necessary for the concept of plurality, while a set size of seven represents a much more maximal example of plurality. The underlying assumption in the complexity approach to plurality concepts is that a set size of two items would be inherently more difficult to judge compared to a set size of seven. Therefore, a research paradigm designed to examining the effects of training plural with larger set sizes would train large vs. small set sizes and examine participants’ performance on untrained items of small and large set sizes. The prediction based on the complexity approach is that training small set sizes would not
generalize well to large set sizes, but training large set sizes would generalize well to large and small set sizes.


