THE MENTAL REPRESENTATIONS OF ALLOMORPHS:

AN INVESTIGATION WITH ARTIFICIAL GRAMMAR LEARNING

by

Hyun Jin Hwangbo

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

Summer 2018

THE MENTAL REPRESENTATIONS OF ALLOMORPHS:

AN INVESTIGATION WITH ARTIFICIAL GRAMMAR LEARNING

by

Hyun Jin Hwangbo

Approved: _____
Benjamin Bruening, Ph.D.
Chair of the Department of Linguistics and Cognitive Science

Approved: _____
George Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____
Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Jeffrey Heinz, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

William Idsardi, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Benjamin Parrell, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Irene Vogel, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Chapter

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation studies the nature of lexical representation by conducting behavioral experiments in the artificial grammar learning paradigm.

It begins with an important question in the phonological and morphological traditions: How are the pronunciations of morphemes represented in long-term memory? When there are multiple pronunciations of a morpheme (e.g., the English plural [-s], [-z], [-ɪz]), is each of the forms stored separately, or is only one form stored, with the other pronunciations derived by phonological transformations? In this dissertation, I refer to the former view as Morpheme Alternant Theory (MAT) and the latter view as Generative Phonology (GP), following Kenstowicz and Kisseberth (1979).

This question is fundamental to phonology because the role phonology plays in grammar is significantly different according to each view. MAT argues that phonology plays a minimal role, and thus predicts that the pronunciations of the allomorphs are memorized or listed as their own forms as the underlying representations. GP argues that phonology is crucial in organizing the mental lexicon and in language, and thus phonology has greater priority than morphological listing.

This dissertation brings a new kind of evidence to bear on the question. Artificial grammar learning experiments were conducted which compared how well learners could learn allomorphy patterns that were more or less amenable to a phonological analysis under the GP view. The results consistently reject the view that speakers just memorize which allomorphs are used in which phonological environments. A more careful examination of the results also shows they are inconsistent with the views that the subjects either prioritized phonological rule learning or that they infer phonological rules only after memorizing which allomorphs are used in which phonological

environments. Instead, the results point to a more complex interplay between phonological and morphological learning. These results provide a foundation and direction for future research. Behavioral data can provide further evidence for how the mental representations of allomorphs are structured in our minds.

Chapter 1

INTRODUCTION

An important question in language science is how the pronunciations of morphemes are represented in long-term memory. This question is particularly acute when one considers that many morphemes within a language have multiple pronunciations. For example, the English plural suffix is pronounced in three different ways, [-s], [-z], and [-ɪz], and the Korean nominative case marker is pronounced in two different ways, [-i] and [-ka]. Is each of these pronunciations stored in long-term memory? Or is only one form stored, from which the various pronunciations are derived? These are long-standing questions in linguistic theory that have been subject to much research in the phonological and morphological traditions.

We call the forms we store in long-term memory *underlying representations* (UR). The underlying representations are important because they allow one *to capture phonological generalizations in a language, and the knowledge of its speakers*, so that "what is in the language" and "what is in the head" of a speaker can be explained (Hyman, 2018, p.7). Hyman argues that underlying representations must have four qualities (p.14) as listed in (1.1).

(1.1)   a. Simple: The URs are not complicated or unnecessarily abstract.

       b. Efficient: The URs describe the phenomena in parsimonious terms.

       c. Restricted: The URs do no more than what can be expected of them to do.

       d. Motivated: The URs are posited to capture productive alternations across words.

As Scobbie et al. (1996) point out, "The general consensus has been that underlying and surface representations are both necessary – that they are distinct. At the same time, the general consensus has been that they are minimally distinct (p.697)." This is where the old but fundamental question comes in: is there a single underlying representation for the alternants, or are there multiple underlying representations in our mind?

Positing a single underlying representation for the pronunciations of morphemes meets these four qualities. There are additional theoretical reasons to prefer a single underlying representation to multiple representations. A single underlying representation in conjunction with phonological transformations posits a *simple and elegant* solution (Odden, 2014, p.80), thus meeting the standards of *Occam's razor*, also known as the Principle of Parsimony (Sober, 2015). Parsimonious theories allow for "explanatory and predictive power" (Anderson, 1985, p.312), and 'predictability' is one of the most important criteria for deciding upon an underlying representation (Hyman, 1975). Languages show systematic variation in the phonology of different morphemes. For example, in English, the possessive and third person singular verbal inflection suffixes are pronounced in the same three different ways as the plural suffix. By positing a single underlying representation, resulting in multiple pronunciations of morphemes after undergoing phonological transformations, one can predict that other morphemes with similar pronunciations would behave the same way. This systematically captures phonological generalizations and the knowledge of the speakers. Therefore, the systematicity of a language is better predicted and captured if there is a single underlying representation and a phonological grammar. I refer to this theory as Generative Phonology (hereafter, GP). Most generative phonologists argue that GP also allows morphemes to have multiple underlying representations—but *only* when it is necessary. The preferred analysis is to posit a single underlying representation.

However, there are researchers that argue that every pronounced variant is stored as an underlying representation. I refer to the theory arguing for multiple underlying representations as Morpheme Alternant Theory (hereafter, MAT), following

Kenstowicz and Kisseberth (1979). There are countervailing reasons to prefer multiple underlying representations to a single underlying representation. Positing multiple underlying representations obviates the need for abstraction because the underlying representations are the observed phonetic forms. Some proponents of this view today are represented by the sublexical approach (Becker and Gouskova, 2016). The sublexical approach (Allen and Becker, 2015) states that the model "does not posit an underlying representation" (p. 14) and thus there is "no search for underlying representations and no further manipulation" (p. 2). This perspective is shared by other researchers in other traditions. For instance, Declarative Phonology (Scobbie et al., 1996) also states that an underlying representation is "a set of distinct surface representations" (p.697). In other words, the surface representations are not derived from another but are selected as "the rules of selection directly state the context in which each alternant may occur. In particular, then, there is no notion of an underlying form from which all surface alternants are derived" (Kenstowicz and Kisseberth, 1979, p.181). In these views, surface representations are not different from underlying representations, as all the alternants are listed, and thus, there are no phonological transformations.

Both the GP view and MAT view are discussed in more detail in §2.1.

The goal of this thesis is to bring a new kind of exprimental evidence to bear on this question. A series of experiments was conducted to study whether either of the two views was supported in an artificial grammar learning paradigm. In one condition, called the GP condition, the allomorph[1] pattern was more amenable to a GP analysis. That is, the alternations are based on a phonological transformation, and thus, one can posit a single underlying representation for the alternants. In the other condition, called the MAT condition, allomorphs were designed so they would

---

[1]The notion of allomorphy is used in a broader sense in this dissertation. Allomorphy is defined as pronunciations that involve "more than one lexical representation" (Faust and Lampitelli, 2016, p.230), and thus the different pronunciations derived from a single underlying representation via phonology are not considered as allomorphy. However, I use the term *allomorphy* to indicate phonologically conditioned patterns of variation in the pronunciations of a morpheme.

be best analyzed as multiple underlying representations because the alternations are phonologically unmotivated, and thus the alternants cannot be abstracted to a single underlying representation.

Across all conditions and experiments (Experiments 1A, 1B, and 2), subjects performed significantly better than chance, which means that they learned the patterns. Two models of morpho-phonological learning in the GP tradition were proposed: the P-First model and the M-First model. In the P-First model, phonological learning is activated before morphological learning, and in the M-First model morphological learning is activated before phonological learning. The MAT-view and the M-First model predict no difference in learning a new morpho-phonological pattern between the two conditions and the P-First model predicts subjects will respond differently in the two conditions. In general, the results showed a difference in learning between the two conditions and thus they reject the MAT-view and the M-First model consistently. However, the results also show the direction of learning in the two conditions does not favor the the P-First model as expected, which points to a more complex interplay in the learning of phonology and morphology.

The most interesting results were from Experiment 2, in which the number of allomorphs to learn was increased. While the results were the same, the rate of selecting the grammatical suffix dropped significantly in the MAT condition compared to Experiments 1A and 1B. No such drop was observed in the GP condition. These results suggest that people have more difficulty learning patterns that have less predictability or systematicity. These results raise questions about the role of phonology when learning a language and how much information is needed for people to infer the phonology of a new language.

While the results do not definitely answer the fundamental question laid out at the beginning of this chapter, they are shown to be robust and replicable. This dissertation brings new evidence to bear on the fundamental question, and shows that there is more than just morphological listing when learning a new language through

4

behavioral experiments that tasked subjects to learn to identify the correct pronunciation of a suffix in an artificial language. This dissertation also provides a foundation and direction for future experiments that would address these questions.

The structure of this dissertation is as follows. Chapter 2 discusses GP and MAT in more detail and provides linguistic evidence for each theory. It also reviews acquisition and psycholinguistic studies, as well as research that used the artificial grammar learning paradigm for learning phonology. Chapter 3 raises the research question, introduces the artificial grammar for this dissertation, and discusses the problem of phonotactics. Chapters 4 and 5 present and discuss experimental evidence. Chapter 4 presents the first set of the experiments, where subjects learned one pair of suffixes, and Chapter 5 presents the second experiment, where subjects learned two pairs of suffixes. Chapter 6 provides general discussion and, lastly, Chapter 7 provides a conclusion.

Chapter 2

BACKGROUND

In this chapter, I will broadly introduce the two theories, GP and MAT, in §2.1, and then present the language data bearing on those theories in §2.2. This section includes language data that favors GP – Kinyarwanda (§2.2.1), Klamath (§2.2.2), and Aghem (§2.2.3) – as well as data that favors MAT – Korean (§2.2.4) and Romanian (§2.2.5). I next review related behavioral studies in §2.4. These behavioral studies address the acquisition of Korean (§2.4.1), experimental psycholinguistic studies of the English plural suffix (§2.4.2), and regular/irregular past tenses (§2.4.3). In §2.5, I review studies of learning phonological patterns in artificial grammar learning paradigms. Lastly, §2.6 summarizes the chapter.

2.1 Generative Phonology (GP) and Morpheme Alternant Theory (MAT)

In this section, I briefly introduce the two opposing theories in broad terms. GP and MAT mainly differ in the mental representations they assign to morphemes (Kenstowicz and Kisseberth, 1979). GP argues that there is a single underlying representation for the morpheme alternants while MAT proposes there are multiple underlying representations. More precisely, in GP, typically there is a single underlying representation of a morpheme in the mental lexicon, and the various surface forms are derived via phonological transformations. In MAT, each of the alternants is listed in the mental lexicon, and the correct alternant is selected by a set of phonologically-conditioned word-formation rules.

Consider one familiar example, the English plural suffix *-s*, which has three alternants, [-s, -z, -ɪz]. This suffix is pronounced differently depending on its phonological environment. It is pronounced as [s] after voiceless non-sibilants (e.g., *books, cats*), as

[z] after voiced non-sibilants (e.g., *boys, chairs, dogs*), and as [ɪz] after sibilants (e.g., *wishes, buses*). The question for theory is: How is the English plural suffix represented in the lexicon? There are two possibilities: abstract a single underlying representation from the alternants or list all surface representations as the underlying representations.

In GP, only one underlying representation, /-z/, is represented. It is the same form as one of the surface representations but shows some differences from the other two surface representations. The alternants, [-s] and [-ɪz], are derived from it by phonological transformations such as devoicing and vowel epenthesis. What matters here is not whether the phonological transformations are described with Rule-based phonology (Chomsky and Halle, 1968) or with Optimality Theory (Prince and Smolensky, 1993), but that there is a lawful transformation from underlying representations to surface representations.[1]

This transformation accounts for the various pronunciations of the suffix. Figure 2.1 schematizes the view of GP and the process with respect to the English plural.[2] In the lexicon, abstract morpheme pronunciations are paired with their meaning. Morphological word formation concatenates morphemes; even after concatenation all three lexical items have the same underlying representation for the plural suffix. The concatenation of the lexical items and the suffix go through phonological transformations, which result in the phonetic forms that are articulated.

---

[1]In the rest of the dissertation, I will use rewrite-rule style notations for the transformations because they are succinct and familiar.

[2]This schema is a simplified version and does not consider the details of more sophisticated models such as Lexical Phonology (Kiparsky, 1982).
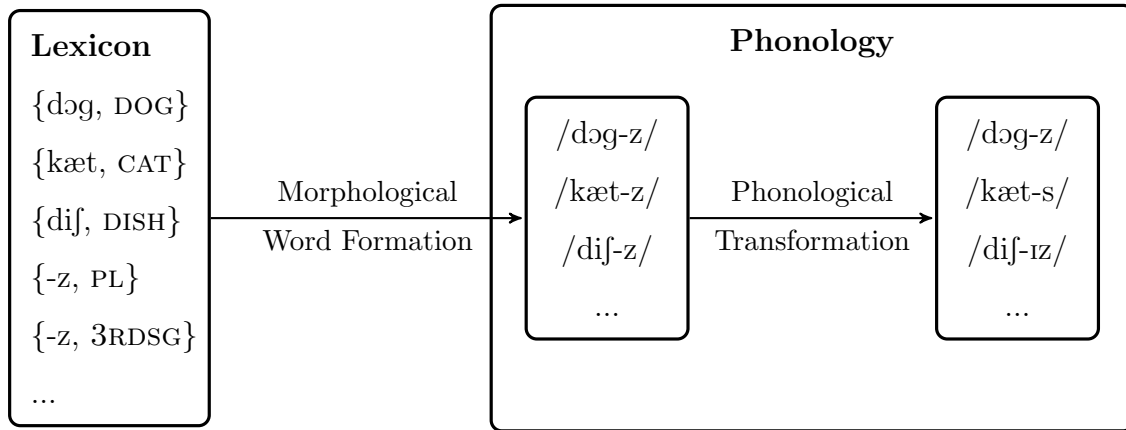
Figure 2.1: Schema of the GP process for underlying representations of morpheme alternants and their realization

Having a single underlying representation for the English plural allows GP to explain and generalize the variations in a systematic and predictable way. In other words, similar underlying representations will be affected similarly by the same phonological processes. For example, English plural alternants are similar to English verbal inflection, third person singular verb alternants. The morpheme is pronounced as [-s] after voiceless non-sibilants (e.g., *walks*), as [-z] after voiced non-sibilants (e.g., *jogs*), and as [-ɪz] after sibilants (e.g., *washes*). The fact that the alternants are the same can be understood when the underlying morphemes are /-z/ and are subject to the same phonological transformation. In other words, in GP, the surface representations and the underlying representations may be different to some degree but having a single underlying representation with phonological transformations explains this systematic variation.

However, this does not mean that GP only argues for a single underlying representation for all of the cases. There are cases where one must list multiple underlying representations. In other words, GP allows both single representation and multiple representations of morpheme alternants. Then the question is when GP posits a single underlying representation with phonology and when GP posits multiple underlying

8

representations with selection rules. At a minimum, GP posits a single underlying representation when morphemes' allomorphs exhibit systematic phonological variation. If the morpheme alternation is arbitrary — if it does not pattern with the phonological generalization of the language — only then are multiple underlying representations listed. The examples of this case are discussed in §2.2.4 and §2.2.5.

Unlike GP, MAT posits that all three alternants, /-s/, /-z/, and /-ɪz/, are listed in the lexicon. There is no difference between the surface representation and the underlying representation, and thus, there is no phonological module and phonological transformations in this view. By listing the surface representations as the underlying representations, the two representations are "minimally distinct (p.697)" (Scobbie et al., 1996).

Since there is no need to abstract a single underlying representation, it is more like classifying the words into different categories – the ones that end with a voiceless non-sibilant obstruent, a voiced non-sibilant obstruent, and a sibilant – so that each of the morphemes can be attached to the right class of words. In sublexical analysis (Allen and Becker, 2015; Becker and Gouskova, 2016), each of the categories are called a sublexicon, and each sublexicon has its own phonotactic rules. Likewise, if the surface representations are the same as the underlying representations, one needs to find selection rules for each category.

The pronounced alternant is selected by phonologically-conditioned word-selection rules. The phonologically-conditioned word-selection rules for the English plural are in (2.1).

(2.1) English plural suffix selection rules:

    a. Select the alternant /-s/ after a word ends with a voiceless non-sibilant obstruent.

    b. Select the alternant /-z/ after a word ends with a voiced non-sibilant obstruent.

    c. Select the alternant /-ɪz/ after a word ends with a sibilant.

One of the main differences between GP and MAT is that the rules described in (2.1) apply only for the plural suffix. In other words, each suffix has its own phonologically-conditioned word-selection rules. In other words, it is more like having multiple grammars for each subclass of words (e.g., sublexical analysis (Allen and Becker, 2015; Becker and Gouskova, 2016) and Cophonology Theory (Inkelas, 2014)). Thus, there must be a separate, though very similar, set of rules for the third person singular verb suffix. In this way, MAT fails to generalize across phonologically similar morphemes of a language, for which it has been criticized (Kenstowicz and Kisseberth, 1979). In other words, in this view, surface representations and underlying representations show no distinction, but fail to generalize the phonology of the language.

The schema of MAT is represented in Figure 2.2. In MAT, all the suffix alternants are listed in the lexicon. Each of the morpheme alternants is selected based on the rules in (2.1), and then realized in its phonetic form, which is the same as the underlying representation. Thus, there is no phonological transformation in this model.
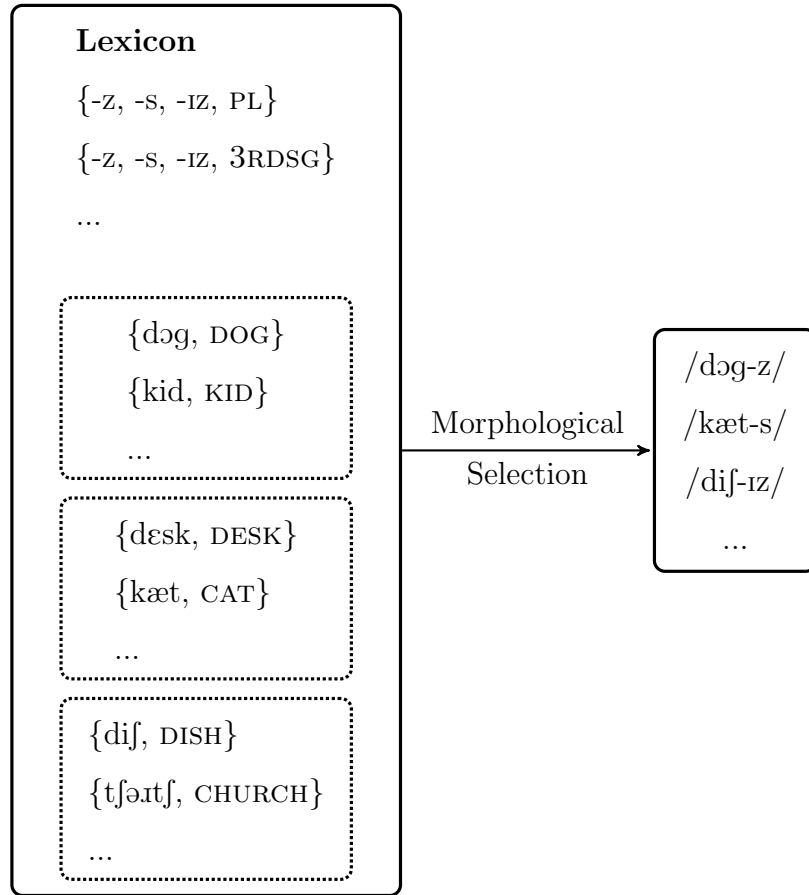
Figure 2.2: Schema of the MAT process for underlying representations of morpheme alternants and their realization

The main difference between GP and MAT lies in whether a single underlying representation or multiple underlying representations is posited. GP has a phonological module where the phonological transformation occurs, whereas MAT does not have a phonological module for phonological transformation. Rather, in the MAT lexicon, the words are categorized based on the suffix selection. Again, it does not mean that proponents of GP only argue for a single underlying representation for all of the cases. Multiple underlying representations are allowed *only* when necessary. But in the MAT-view, all of the alternants are listed on their own.

Many believe phonology is part of the language faculty (Berent, 2013; Hauser

et al., 2002; Fitch et al., 2005; Odden, 2014; Kenstowicz and Kisseberth, 1979, and many other studies). This belief lines up with the four qualities of underlying representations mentioned in (1.1). One can use the phonological module to provide an elegant explanation for the systematicity of sound patterns in languages. Phonology allows the underlying representations to meet all of the four qualities: simple, efficient, restricted, and motivated, as well as predictive. The question is about how much of a role is assigned to phonology compared to morphology.

In this section, I reviewed the two theories in broad terms. GP posits a single underlying representation for allomorphs, while MAT posits multiple underlying representations for allomorphs. This dissimilarity leads to different roles for phonology: it accounts for phonological transformations in GP while playing a minimal role in MAT. In the next section, I review the corpus-internal evidence arguing for each theory in detail.

2.2 Corpus-Internal evidence: Language Data

In this section, I review corpus-internal evidence for each theory. Proponents of MAT believe that all of the underlying representations are the same as the phonetic forms without exceptions. On the other hand, many people from generative linguistics believe that most of the alternants in natural language can be analyzed with a single underlying representation and some alternants that must be analyzed with multiple underlying representations. That is, GP incorporates both. The main purpose of this section is to show that we need both morphological listing and phonology. The first two sections (§2.2.1 and §2.2.2) provide evidence for phonology. §2.2.3 shows a case where the underlying tone must be different from the surface tone, which also supports phonology. §2.2.4 and §2.2.5 provide evidence for having multiple underlying representations for allomorphs.

2.2.1 Kinyarwanda

In Kinyarwanda, the verb root form alternates between [t] ∼ [h] (Kenstowicz and Kisseberth, 1979, pp.183∼184). The verb *tem* 'cut' alternates between *tem* ∼ *hem* where [h] occurs when there is a nasal consonant prefix. Examples are shown in (2.2).

(2.2) a. **t**ema 'cut!'

   b. n**h**emera 'cut for me!'

   c. mu**t**emera 'cut for him!'

   d. ba**t**emera 'cut for them!'

Following GP, the underlying representation of [tem] would be listed as /tem/, as in (2.3a), and undergo phonological transformation, as in (2.3b).

(2.3) a. {tem, CUT}

   b. t → h / n ____

Following MAT, both alternants would be listed as underlying representations, as in (2.4a) and have selection rules, as in (2.4b).

(2.4) a. {tem, hem, CUT}

   b. i. Select /hem/ when it follows a nasal consonant.

     ii. Select /tem/ otherwise.

Now consider the examples in (2.5) of more words in which [t] and [h] alternate. Following GP, the phonological transformation in (2.3b) can explain all these words and their alternants, so there is no need to add more phonological transformation rules. Following MAT, the words and the alternants are listed in the lexicon as in (2.6), with the selection rules as in (2.7), because the rules in (2.4b) only apply to that lexical item.

(2.5) a. **t**uma 'send!' n**h**uma 'send me!'

   b. **t**eeka 'cook!' n**h**eekera 'cook for me!'

   c. **t**egereza 'wait!' n**h**egereza 'wait for me!'

(2.6) a. {tum, hum, SEND}

　　　　b. {teek, heek, COOK}

　　　　c. {tegerez, hegerez, WAIT}

(2.7) a. i. Select /hum/ when it follows a nasal consonant.

　　　　　　ii. Select /tum/ otherwise.

　　　　b. i. Select /heek/ when it follows a nasal consonant.

　　　　　　ii. Select /teek/ otherwise.

　　　　c. i. Select /hegerez/ when it follows a nasal consonant.

　　　　　　ii. Select /tegerez/ otherwise.

The rules in (2.4b) and (2.7) are basically referring to the same phonological alternation in the context of [t] and [h]. The main problem with the rules in (2.4b) and (2.7) is that they lack predictability and efficiency. The rules are not efficient because the same phenomenon is described by separate rules for each lexical item. In contrast, GP can efficiently explain and capture the predictable alternations across the words.

　　　If we relax the MAT-style rules, they may be combined into one as in (2.8).

(2.8) "If a morpheme has two alternants, one beginning with /t/ and the other beginning with /h/, use the alternant with /h/ after a nasal consonant and use the other alternant elsewhere." (Kenstowicz and Kisseberth, 1979, p.184)

Selection rule (2.8) seems to solve the efficiency problem, but the predictability problem still remains: the selection rule only works for the words that meet the conditions of the rule, and thus lacks the predictability for other words. For example, if there is a new word that begins with /t/, the rule (2.8) does not predict that the word will begin with [h] after the nasal, because the rule applies only *if* there are two alternants. For a new lexical item, however, one does not know whether there are two alternants or not. Moreover, nothing prevents the lexicon from including /t/-initial lexical items that does not alternate. Therefore, the predictability problem still holds.

As mentioned before, a linguistic generalization must capture a property of the language systematically, and predict the patterns of the language. With enough data, the alternation of [t] and [h] is predictable in Kinyarwanda. Thus, the rule (2.3b), the GP-style approach, predicts the behavior of the language better, and is a more succinct and efficient way to describe the language. Because the phonological transformation is predictable using the GP-style analysis, positing a single underlying representation for /t/ with a transformation rule that changes to [h] after a nasal consonant is a stronger and more scientific analysis of the language than listing all the alternants with their own selection rules.

2.2.2   Klamath

A complicated or an opaque transformation of the prefix alternants and the verb stems clearly is difficult for MAT (Kenstowicz and Kisseberth, 1979). Consider examples of the Klamath causative prefix, which alternates among [sne] ∼ [sno] ∼ [sna] as in (2.9).

(2.9)	a.	ge̱ːǰig-a	'is tired'	sn**e**-ge̱ːǰig-a	'makes tired'
	b.	qdoːč̌ -a	'it rains'	sn**o**-qdoːč̌-a	'makes it rain'
	c.	m'aːsʔ-a	'is sick'	sn**a**-m'aːsʔ-a	'makes sick'

The prefix alternates the vowel to match the vowel in the stem. Following the MAT-style, all three alternants are listed as the causative prefix: {*sne-, sno-, sna-,* CAUSATIVE}, and the selection rule for the prefix is in (2.10).

(2.10)	Select the alternant of the causative prefix that has the same vowel as the following verb.

Rule (2.10) seems to describe the language well in terms of efficiency and predictability. Now consider the following examples in (2.11), where the vowels of the stem are deleted when the stems occur with the causative prefix.

(2.11)	a.	pag-a	'barks'	sna-**pg**-a	'makes a dog bark'

b.	nqot'-a	'scorches'	sno-**nqt'**-a	'scorches s.t.'
c.	wet-a	'laughs'	sne-**wt**-a	'makes laugh'

Verbs with a short vowel have alternants: {*pag, pg,* BARK}, {*nqot', nqt',* SCORCH}, *and* {*wet, wt,* LAUGH}. The alternants without a vowel is selected after the causative prefix, therefore, we can generalize the selection rule for the verbs as in (2.12).

(2.12)   Select an alternant of the verb that has no vowel after the causative prefix.

Given the two selection rules (2.10) and (2.12), MAT-style analysis has difficulty forming the words in (2.11) because we do not know which rule to apply first. The concatenation of the causative prefix and the verb is illustrated in (2.13).

(2.13)   CAUSATIVE   VERB

$$
\left\{
\begin{array}{c}
\text{sne-} \\
\text{sno-} \\
\text{sna-}
\end{array}
\right\}
\; + \;
\left\{
\begin{array}{c}
\text{pag} \\
\text{pg}
\end{array}
\right\}
\; + \; \text{a}
$$

There are two possible ways to explain the concatenation: either apply selection rule (2.10) before rule (2.12), or vice versa. If we apply selection rule (2.10) first, we cannot choose the alternant of the causative prefix unless it refers to the verb form *pag*. However, this contradicts MAT, which argues that the underlying form and the surface form are the same. Since the form *pag* is not in the surface form with the causative prefix as shown in (2.11a), the rule (2.10) cannot be applied first. If we apply selection rule (2.12) before rule (2.10), the alternant *pg* will be selected because it follows the causative prefix. However, MAT now has a problem in selecting the causative prefix because the verb does not have a vowel. Regardless of which rule you apply first, the MAT-style analysis contradicts itself and does not capture the phonological transformation of the language well.

If we follow the GP-style, the morphology is explained without contradiction and it predicts the patterns of the morphemes. Briefly, the words in (2.9) and (2.11) can be explained by transformations affecting the underlying vowel in the prefix and deleting

16

the short vowel in the stem, and the two transformations occur in that order. During the vowel agreement transformation, the prefix vowel is set to agree with the vowel in the verb stem. Next, the vowel in the verb stem is deleted. This transformation allows us to have a grammatical surface representation. An example of this transformation is shown in (2.14).

(2.14) /snV + pag + a/   Underlying Representation
           sna + pag + a   Vowel agreement
           sna + pg + a   Short vowel deletion (in the stem)
         [sna-pg-a]   Surface Representation

By having a single underlying representation and the transformations, the GP-style analysis predicts that a novel verb will undergo the same transformation, and thus conforms to the language pattern. For such cases, GP explains the language more efficiently and predictably than MAT.

2.2.3   Aghem

The MAT-style analysis posits that the surface form and the underlying forms are the same, and thus, lists all alternants in the lexicon. In this section, I show a case where the two words have the same surface tones but the underlying tones must be different.

Consider Aghem (Hyman, 1979, 2011, 2018). The examples in (2.15) show the prefix kɨ́- with an H (high) tone followed by a stem with H. The tones are indicated by the diacritics, and H stands for high tone, L stands for low tone.

(2.15) a. H-H   kɨ́-fé   'leg'
        b. H-H   kɨ́-wó   'hand'

In both examples, the stem and the prefix are pronounced as H in isolation. With the demonstrative kɨ́n, the noun class prefix kɨ́- deletes, and the demonstrative kɨ́n alternates between [kɨ́n] and [↓kɨ́n] or H and ↓H. (2.16) shows the examples.

17

(2.16) a. H-H   fé kín   'this leg'

         b. H-⁺H   wó ⁺kín   'this hand'

Both stems are H, yet one takes the H and the other takes ⁺H. GP analyzes the two stems as having an underlying difference that causes the alternation of the demonstrative. In other words, the word *fé* is underlyingly H, (2.17a), while the word *wó* is underlyingly HL, (2.17b), where L is the floating tone that causes the downstep of the following H.

(2.17)   a. $\left\{ \begin{array}{ll} \text{fe,} & \text{LEG} \\ \text{H} & \end{array} \right\}$

         b. $\left\{ \begin{array}{ll} \text{wo,} & \text{HAND} \\ \text{HL} & \end{array} \right\}$

Unlike GP, MAT lists both stems with H because the stems do not show alternation. Therefore, MAT will list both allomorphs as {kín, ⁺kín, THIS} with selection rules such as 'select the ⁺H alternant after an H stem'. However, this is problematic for MAT because both stems are H; thus, the stem and suffix cannot be correctly selected as shown in (2.16). The solution for MAT is positing different types of H for the stems. For example, *fé* is Type 1 H and *wó* is Type 2 H. Given this categorization, the selection rules can be 'select the ⁺H alternant after a Type 2 H stem' and 'select the H alternant after a Type 1 H stem.'

Now, consider the following examples, in which the suffix alternates between the falling (HL) and L.

(2.18)   a.   H-HL   fé kîa   'your sg. leg'

         b.   H-L   wó kìa   'your sg. hand'

Following the GP-style analysis, the falling tone in (2.18a) results from a spreading H tone from the stem to the suffix. The L tone in (2.18b) results from floating L, which blocks the spreading. In other words, the alternants are exactly what is expected for the underlying tones and the transformation.

As above, MAT would differentiate the H tone into Type 1 and Type 2 for the selection rules. Words are listed as {kɨ̂a, kɨ̀a, YOUR SG.} in the lexicon, with selection rules such as 'select falling (HL) after a Type 1 stem' and 'select L after a Type 2 stem.' It seems that there is no contradiction for MAT in this analysis as it can accommodate the data. However, MAT cannot explain the language economically and parsimoniously (Sober, 2015), nor can MAT predict the patterns of the language. GP more efficiently captures the behavior of the language and can predict its patterns.

In summary, Aghem shows that even when the surface tones are the same, the underlying tones can be different. Positing different underlying tones, as in the GP-style analysis, allows one to analyze the language efficiently and predict the patterns of the language. Positing the same surface form and underlying representation, as in the MAT-style analysis, fails to analyze the language efficiently or predict the observed patterns.

2.2.4 Korean

The language data in the previous sections showed that it is more efficient and predictable to posit a single underlying representation than having multiple representations for alternants. We also reviewed a case where the surface forms have the same tone, but the underlying tones are different.

However, there are other cases in which the alternants need to be listed. One of these is the Korean nominative case marker, which alternates between [-i]∼[-ka], as shown in (2.19) (Sohn, 1999).

(2.19) a. pal-i   'foot-NOM'
        b. kʰo-ka   'nose-NOM'

A word takes the nominative case marker [-i] when it ends with a consonant, and takes [-ka] when it ends with a vowel. Under the GP-style analysis, the underlying representation of the alternants should be the same as or similar to one of the alternants to posit a single underlying representation, and the phonological transformation results

in the appropriate surface representation. So there are three possible ways to posit an underlying representation: /-i/, /-ka/, or /-ika/. If the underlying representation is /-i/, then there must be two transformations: vowel lowering and consonant epenthesis. The vowel lowering transforms the vowel /i/ to /a/, and the consonant epenthesis transforms the consonant /k/ to be epenthesized. However, there is no phonological motivation for lowering the vowel or epenthesizing the consonant. Moreover, these transformations have no other place in Korean phonology. Therefore, positing the underlying representation as /-i/ is not favored. If the underlying representation is /-ka/, then the consonant must be deleted and the vowel must be raised when a word ends with a consonant. Again, there is no phonological motivation for these processes, so positing the underlying representation as /-ka/ has the same problem as positing the underlying representation as /-i/. Lastly, if the underlying representation is /-ika/, then the high front vowel must be deleted when a word ends with a vowel to surface as [-ka], and the consonant and the low vowel must be deleted when a word ends with a consonant to surface as [-i]. Again, the processes have no phonological motivation; therefore, positing the underlying representation as /-ika/ is not favored either. Since the two alternants [-i] and [-ka] are not phonologically related, it is difficult to predict which one is the underlying representation, and how the transformations occur.

In addition, the alternation is arbitrary and not motivated in Korean. If the alternation is intended to avoid a consonant cluster, then there should be no consonant followed by another consonant in Korean. This is not the case: Korean allows a consonant to be followed by another consonant and a vowel to be followed by another vowel (e.g., [dʌp-ko] 'cover-and', [t͡ʃa-e] 'at the car'). This is the opposite of the nominative case marker, which does not allow a consonant followed by a consonant or a vowel followed by a vowel. Therefore, the GP-style analysis of the nominative case marker does not capture the linguistic behavior of Korean. The transformations that would have to be posited are ad-hoc. This discussion adds to the point that GP also needs morpheme listings for some alternants.

The Korean nominative case marker is more amenable to the MAT-style analysis. The alternants would be listed as nominative markers: {*-ka, -i*, NOM}, and have phonologically-conditioned selection rules, as in (2.20) (Nevins, 2011).

(2.20)   Korean nominative case marker selection rules:

    a. Select /-i/ as a nominative case marker when a word ends with a consonant.

    b. Select /-ka/ as a nominative case marker when a word ends with a vowel.

As shown in (2.21), whenever the nominative case marker occurs after a noun, one of the alternants is selected according to the selection rules of (2.20). This is a much simpler way to explain the pattern. Because this selection rule only applies to the nominative case marker, a consonant followed by another consonant would not be a problem for other markers.

(2.21)   NOUN   NOM

$$\{\text{pal}\} \quad + \quad \begin{Bmatrix} \text{-i} \\ \text{-ka} \end{Bmatrix}$$

Following the MAT-style analysis, listing all the alternants, the alternants are selected according to the rules rather than being derived from a common underlying representation. Because the alternants are phonologically unrelated and the selection does not reflect any phonological phenomena in Korean, listing all the alternants with phonologically conditioned selection rules is a more plausible analysis than positing a single underlying representation with phonological transformations for the nominative case marker.

2.2.5   Romanian

Examples from Romanian (Steriade, 2008; Nevins, 2011) also favor the MAT-style analysis. Romanian shows k-palatalization before front vowels and glides. This alternation is shown in (2.22) for the plural suffix [-i̯]; all examples in this section are from Steriade (2008).

(2.22)   a.   sʌrák-(u)   'poor-MASC.SG'

           sʌrátʃ-i̯   'poor-MASC.PL'

      b.   núk-ʌ   'nut-SG'

           nútʃ-i̯   'nut-PL'

Following the GP-style analysis, the transformation can be represented as in (2.23). The voiced counterpart [g] also alternates with [dʒ], but for simplicity I only represent the voiceless.

(2.23)   k → tʃ / ＿＿ [−back, +vocalic]

If Romanian has such a phonological transformation, the /k/ sound should always transform to [tʃ] before front vowels.

Now, consider the examples in (2.24). These nouns end with *k* and occur with the plural suffix [-uri̯] rather than [-i̯].

(2.24)   a.   lók-(u)   'place-SG'

           lók-uri̯   'place-PL'   *lótʃ-i̯

      b.   fok   'fire-SG'

           fók-uri̯   'fire-PL'   *fótʃ-i̯

The GP-style analysis has difficulty with the alternation of the plural suffix because the alternation between /-uri̯/ and /-i̯/ is arbitrary. One could posit a single underlying representation of the suffix as /-uri̯/ and delete the high back vowel and the rhotic after /k/, but there is no phonological motivation for the deletion. In addition, /-uri̯/ can come after /k/ as well. Instead, one could posit two classes of nouns and the underlying representation of the plural suffix as /-uri̯/, with the vowel and consonant deletion after a certain class of nouns. However, this would still lack a phonological motivation or explanation for the deletion. Therefore, positing a single underlying representation seems like an unnecessary abstraction.

Instead, one can posit multiple underlying representations with selection rules. The plural suffix [-i̯] occurs only when the stem has an allomorph that ends with

22

alveopalatal [tʃ]. The words in (2.24) show no palatalization, and thus, do not oc-
cur with the suffix [-i]. Therefore, the plural suffix alternation between [-i]∼[-uri] is
captured. This can be stated in plural suffix selection rules as in (2.25).

(2.25)   Plural suffix selection rules:

    a. Select /i̯/ when a noun stem has an allomorph ending with alveopalatal
       /tʃ/.

    b. Select /uri̯/ otherwise.

In addition to the plural suffix, derived verb endings also show the alternation
depending on the stem. That is, stems that show alternation with the plural suffix
take the verb ending [í] as in (2.26a), whereas the stems that do not show alternation
with the plural suffix take the verb ending [á] as in (2.26b). Again, the alternation of
the verb endings [í]∼[á] is difficult to generalize following the GP-style analysis, which
would also lack phonological motivations.

(2.26)   a.   kolák   'bagel-SG'

          kolátʃ-i̯   'bagel-PL'

          i̯ŋ-kolátʃ-í   'to roll up'

    b.   fok   'fire-SG'

          fók-uri̯   'fire-PL'

          i̯n-fok-á   'to fire up'   *i̯n-fok-í

Following the MAT-style analysis, the verbal suffix selection rule can be stated
as in (2.27).

(2.27)   Verbal suffix selection rules:

    a. Select /í/ when a noun stem has an allomorph ending with alveopalatal
       /tʃ/.

    b. Select /á/ otherwise.

The alternants of the stem words and the suffixes are all listed in the lexicon as,
for example, {kolák, kolátʃ, BAGEL}, {í, á, VERBAL SUFFIX}, and {-i̯, -uri̯, PLURAL}.

A word without an alternant would be listed by itself. The schema of lexical items are presented in (2.28) under the rules in (2.25) and (2.27).

(2.28)

a.   BAGEL   PL   b.   BAGEL   VERBAL SUFFIX

$$\left\{ \begin{array}{c} \text{kolák} \\ \text{kolátʃ} \end{array} \right\} + \left\{ \begin{array}{c} \text{i̯} \\ \text{uri̯} \end{array} \right\} \qquad \left\{ \begin{array}{c} \text{kolák} \\ \text{kolátʃ} \end{array} \right\} + \left\{ \begin{array}{c} \text{í} \\ \text{á} \end{array} \right\}$$

c.   FIRE   PL   d.   FIRE   VERBAL SUFFIX

$$\{\text{fok}\} + \left\{ \begin{array}{c} \text{i̯} \\ \text{uri̯} \end{array} \right\} \qquad \{\text{fok}\} + \left\{ \begin{array}{c} \text{í} \\ \text{á} \end{array} \right\}$$

Because all the alternants of the stem and the suffixes are listed, there is no need to classify the stems that undergo transformation and those that do not. The appropriate alternant is selected according to the selection rules of (2.25) and (2.27). When there is an alternant that ends with /tʃ/, it will be always the case that /i̯/ is selected as a plural and /í/ as the verbal suffix. Therefore, Romanian k-palatalization and suffix alternations are better explained by the MAT-style analysis.

The Romanian examples show that the MAT-style analysis is more suitable than the GP-style analysis. The GP-style analysis lacks phonological motivation for the transformation, as was the case for the Korean nominative case marker. It is important to point out that the MAT-style analysis is more suitable only when the GP-style analysis fails to be motivated.

In conclusion, this section, §2.2, shows that phonology of most language data is best captured when positing a single underlying representation, but there are also data that are best explained when listing multiple underlying representations. That is, rather than listing all of the alternants, languages are best generalized by having both phonology and morphological listing, as GP argues. It is important to determine which analysis captures the behavior of a language in a simpler and more efficient

way. Regarding efficiency, most of the linguistic data favor the GP-style analysis by positing a single underlying representation for the alternants. It is also important to consider predictability. As described in the above sections, the GP-style analysis is able to predict and therefore explain the allomorphy of language, whereas the MAT-style analysis lacks the ability to do so. This is the weakness of listing all the alternants with selection rules. However, there are some cases where listing all the alternants with selection rules better captures the alternation. In these cases, morphological listing, the MAT-style analysis, better explains the pattern.

2.3 Phonology or Morpheme Listing?

Broadly, there are two views on underlying representations as we have seen in § 2.1. One of the main differences between the two views is when a theory allows multiple morpheme listing. In the view of GP, multiple listing is *only* when it is necessary, and otherwise, the alternants have a single underlying representation. The surface representations are results of phonological transformations. The main reason why people argue for this view is that positing a single underlying representation with phonological derivations allows one to capture phonological behaviors of a language, and thus, can predict morpho-phonological patterns for novel words.

In the view of MAT, all of the alternants are listed in our mental lexicon with selection rules, in other words, there is no phonological derivation. Of course, one can describe a language with this view. But, one cannot generalize phonological behaviors if we just describe what we see and thus MAT fails to predict the pattern of a language.

As we have seen in § 2.2, both views have some level of difficulty capturing all phonological generalizations. In Kinyarwanda (§ 2.2.1), Klamath (§ 2.2.2), and Aghem (§ 2.2.3) MAT shows difficulty, while in the Korean nominative case marker (§ 2.2.4) and Romanian (§ 2.2.5) GP shows difficulty.

The strongest version of GP which argues for a single underlying representation in every case is unsupported. Thus, nowadays, proponents of generative phonology

posit a single underlying representation whenever possible, and posit multiple representations *only* when it is necessary. The reason why they continue to argue for a single underlying representation is because it captures phonological behaviors of a language more systematically than listing all the morphemes.

Next, I discuss behavioral evidence from the acquisition and processing literature regarding the mental lexicon.

2.4 Behavioral Evidence

This section reviews evidence for the acquisition and psycho-linguistic processing of languages that are argued to have a single underlying representation or multiple underlying representations. Behavioral evidence sheds light on theories by reflecting how people treat, process, or learn a language. Since no studies have directly investigated the relationship between underlying representations and their learnability, I review related works. I review acquisition of the Korean nominative case marker, which is argued to have multiple underlying representations, in §2.4.1 and the English plural suffix, which is argued to have a single underlying representation, in §2.4.2. After that, I review the processing of English past tense verbs, which are argued to have a single underlying representation for regular verbs and multiple underlying representations for irregular verbs, in §2.4.3.

2.4.1 Korean nominative case marker acquisition

As reviewed in §2.2.4, the Korean nominative case marker favors the MAT-style analysis. If there are two underlying representations, how do Korean-speaking children acquire the case marker? Do they acquire the forms at the same time or one earlier than the other?

To briefly review, the Korean nominative case marker alternates between *-ka* and *-i*. The former is used with words that end in a vowel (e.g., $k^h o$-*ka* 'nose-NOM') and the latter is used with words that end in a consonant (e.g., $p^h al$-*i* 'arm-NOM').

Korean-speaking children start to produce the nominative case marker between 1;7 and 2;0 (Kim, 1997; No, 2009). Studies have shown that Korean-speaking children acquire -*ka* earlier than -*i* (Kim, 1997; No, 2009; Cho, 1982). The time lag between the acquisition of -*ka* and -*i* is seven to eight months (Kim, 1997).[3] To be more specific, Kim (1997) studied five children longitudinally from 1;7 to 3;5. She reported that all five children acquired the -*ka* form earlier than the -*i* form, at the two-word stage between the age of 1;8 and 2;0. No (2009) studied three children longitudinally, and reported similar results: the nominative case marker -*ka* is produced during the two-word stage from 1;7. Cho (1982) conducted a longitudinal study of three children from age 2;2, 2;7, and 2;10. Two children produced the markers from the beginning of the study, when they were 2;7 and 2;10, and one child produced no nominative or accusative case markers in spontaneous speech. She also reported that the children produced -*ka* more than -*i* in spontaneous speech, and the marker -*ka* was produced correctly 208 times while -*i* was produced correctly 24 times. The errors during the stage are interesting in that all children in the studies made the same kind of error: substituting the markers. That is, the children substituted -*ka* or *-*i-ka* for -*i* in spontaneous speech. Example (2.29) shows that a child produced both errors in one utterance.

(2.29)   *Hammeni,   samchon-*ka thokki manci-e   thokki, samchon-*i-ka*
         grandmother uncle-NOM   rabbit touch-DECL rabbit uncle-NOM
         'Grandma, uncle is touching the rabbit.' (P 2;1)

For the same noun, *samchon* 'uncle', the child produced two errors within an utterance. The child substituted both -*ka* and *-*i-ka* for -*i*.

Example (2.30) shows the latter type of error, substituting *-*i-ka* for -*i*. Note that all the errors involved substituting -*i*; there were no errors that substituted -*ka*. The examples below are from Kim (1997).

---

[3]One of the possibilities for the time lag of the acquisition may be the input frequency. The ratio of -*ka* to -*i* in the input was at least twice and at most eight times higher than the input of -*i* (Kim, 1997).

(2.30)   a. *hyeng-*i-ka   swuyeng   an   hay*
             brother-NOM swimming NEG do:DECL

             'Brother does not swim.' (C 2;3)

         b. *kom-*i-ka   aphu-tay*
             bear-NOM hurt-REPORT

             '(The) bear says (that) (it) hurts.' (P 2;1)

         c. *pholkhein-*i-ka ilehkey   manh-a*
             forklift   truck-NOM so   be.many-DECL

             '(There) are so many forklift trucks.' (C 2;2)

         d. *saykkal-*i-ka ttokkath-ney*
             color-NOM   be.same-FAMILIAR.ASSERTIVE

             '(The) color is the same.'   (J 2;2)

Cho (1982) analyzed the number of errors that the children produced, and the children did not make any errors by substituting *-ka*, but only by substituting *-i*. Out of 39 productions where *-i* was required in spontaneous speech, this error was made 15 times.

These errors indicate that Korean-speaking children seem to learn the phonologically-conditioned rule that it is illegal to use the marker *-ka* with a noun ending in a consonant. Thus, the children might have used **-ika* instead of *-ka* (Kim, 1997). We can speculate that children know the morpho-phonological rule of the Korean nominative case marker because, although children acquired *-ka* earlier, they did not use the form *-ka* for all nouns.

It is interesting to speculate on the possibility that the errors or the time lag between the acquisition of the alternants are due to the need to learn lexical representations. For example, if children posited the underlying representation of the alternant as /-ika/, not /-i/ and /-ka/ separately, and know the morpho-phonological rule that /-ka/ cannot be used words ending in a consonant, they might use the default form /-ika/ so they did not violate the morpho-phonological rule. It is possible to posit /-ika/ as the underlying representation because the alternants occur in the same position

but the markers do not occur at the same time. If children speculate the morpho-phonological rules, they might generalize the process as deleting the vowel after words ending in a vowel. However, as they acquire the language, they realize that such a phonological transformation is not applicable elsewhere in Korean, and so they adjust the underlying representation to /-ka/ and /-i/ separately.

Although it is difficult to find out what Korean-speaking children have in mind regarding underlying representations for the nominative case marker, acquisition of the Korean nominative case marker gives us an interesting insight about allomorphy and its learnability. Setting aside the difference in the input frequency, Korean speaking children acquire the morpho-phonological rule for the nominative case marker that *-ka* occurs with a noun ending in a vowel, and that it cannot come after a noun ending in a consonant early on, and the two alternants are not acquired at the same time.

2.4.2 English plural suffix

The English plural suffix shows systematic alternation, which arguably supports the GP-style analysis. One of the classic studies of the development of allomorphs in an experimental setting is the 'wug-test' (Berko, 1958).

Berko (1958) studied 56 children from age 4 to 7 for their production of English morphemes. Children were asked to produce the plural form (as well as other forms such as the past tense form) of a nonce word. For example, they were introduced a picture that looked like a bird and were told the sentence "This is a wug." After that, they were shown two birds, and they had to finish the sentence "There are two _____." If children know the alternation of the plural suffix, they should produce the [-z] form after a word ending in a voiced obstruent, [-s] after a word ending in a voiceless obstruent, and [-ɪz] after a word ending with a sibilant. The results showed that children correctly produced [-z] and [-s] for nonce words (e.g., *wugs, biks*) from the age of 4. Although the children correctly produced the plural forms of real words with a sibilant ending (e.g., *glasses*), they were not very successful at producing the [-ɪz] for nonce words (e.g., *gutches*). Children showed 91% correct answers for the words

that they already knew, like *glasses*, but only 36% correctness for the nonce words, like *gutches*. The results indicate that children at the age know the plural allomorphs, [-s, -z, ɪz], and are able to extend the phonological rule to new words especially for [-s] and [-z]. It seems that children acquire the morpho-phonological rules for voicing alternation earlier than the vowel epenthesis.

Likewise, the children showed the same pattern with the English possessive suffix *'s*. Children were able to produce the possessive forms of *wug's* and *bik's* with the pronunciations of [-z] and [-s], but the children did poorly on words such as *nizz's*. Again, it seems that children could extend the phonological rule of [-s]∼[-z] alternation to new words, but had difficulty extending the epenthesis rule. In other words, children seem to acquire the devoicing rule earlier than the epenthesis rule.

A similar pattern occurred with regular past tense verbs. Children were able to extend the [-d] and [-t] alternation on the nonce words (e.g., *ricked* and *glinged*)[4]. However, they performed poorly on the [-ɪd] form, such as *motted*. Again, children seem to acquire the devoicing rule earlier than the epenthesis rule.

The study showed that the acquisition of the pronunciations of the morphemes is systematic. The phonological rules or transformations are applied to various morphemes, and thus children show similar patterns in producing the plural, possessive, or past tense forms of the nonce words. That is, they were able to extend the phonological knowledge of voicing alternation of the suffix before they could do so with the vowel epenthesis. If we adopt the MAT-style analysis and have /-s/, /-z/, and /-ɪz/ as the underlying forms of the plural and possessive suffix, there is no reason for children to have difficulty extending the application of /-ɪz/ to nonce words in both plural and possessive forms. If we adopt the GP-style analysis and say children acquire the devoicing rule/transformation before the one for epenthesis, this explains the behavior of children for both the plural and the possessive suffix and also the regular past tense

---

[4]It is possible to produce a irregular form of the word (e.g., *glang* in line with *ring-rang*). However, only one child out of 86 said *glang* and two children said *glanged* which is an irregular form with the regular past tense morpheme (Berko, 1958).

suffix. Therefore, the study provides indirect evidence that the children abstracted the allomorphs to the single underlying representation using phonological transformation rules.

2.4.3 The English past tense: regular verbs and irregular verbs

One of the related debates regarding underlying representations is the English past tense. The English past tense is comprised of two types: regulars, which add the suffix *-ed*, and irregulars, which show suppletive morphology such as *go-went*. Thus, the question of the English past tense comes down to whether the regulars and the irregulars are represented and treated in the same way (e.g., McClelland and Patterson, 2002) or differently (e.g., Pinker, 1998; Pinker and Ullman, 2002). Here, I focus on the latter.

Although the English past tense cases do not directly provide evidence for underlying representations between the alternants in regular verbs, studies show that regulars and irregulars are treated differently. In other words, people treat morpho-phonological transformations differently than the suppletive ones. The Words and Rules theory (Pinker, 1998; Pinker and Ullman, 2002) argues that English past tense verbs are stored in the lexicon in two different ways. Regular past tense verbs are stored in the lexicon with word stems, and the past tense of the verb is derived by the rule: add *-ed* to the verb (e.g., {dʒʌmp, JUMP}, {-d, PAST}). Irregular past tense verbs are stored in the lexicon as they are. For example, the stem word is stored as {ɹʌn, RUN}, and the past tense of the word is stored separately {ɹæn, RUN-PAST}. The argument for separate entries of the stem and the past tense is the irregularity of the English past tense. The English past tense verbs have different forms, and cannot be generalized with phonological transformation rules: *go–went, buy–bought, sing–sang, give–gave*, and so on.

The English irregular past tense verbs are interesting because the irregulars have not only the semantic property of the stem (e.g., RUN), but also the tense feature,

PAST (e.g., {ɹæn, RUN-PAST}). The irregular past tense verbs do not share morpho-phonology with the regular past tense morphemes (adding /-d/, /-t/, or /-ɪd/ to the stem). The irregulars have their own phonological forms. Therefore, if English irregular past tense verbs behave similarly to regular past tense verbs, that indicates that the regular and irregular verbs are represented similarly in the mind, while if irregular past tense verbs do not behave similarly to regular past tense verbs, that indicates that the two verbs are not represented similarly in the mind. Various psycholinguistic studies are related to this topic.

Studies of left-hemisphere impairment patients support the argument that regular and irregular past tense verbs are treated differently (Marslen-Wilson and Tyler, 1997, 1998, 2007; Tyler et al., 2002a,b). In short, the irregular past tense verbs are processed as a whole, while the regular past tense verbs are decomposed into a stem and the past tense morpheme. Tyler et al. (2002b) measured the reaction times and error rates of left-hemisphere deficient patients through a speeded judgment task that used priming words. The study showed that there was no priming effect with the regular past tense verbs, unlike with the irregular past tense verbs. More specifically, the patients were slower with regular past tense words (e.g., *called-call*) compared to irregular past tense words (e.g., *wrote -write*). The slower reaction time with the regular verbs suggests there is a morpho-phonological process that decomposes them into the word stem and the suffix. Irregular past tense verbs did not require a morpho-phonological process. The results indicated that regular and irregular verbs are processed in different ways. Event-related functional magnetic resonance imaging (fMRI) studies support the conclusion that regular verbs are decomposed into stems and morphemes (Marslen-Wilson and Tyler, 2007). Based on the English past tense studies, we can speculate that when forms are phonologically unrelated, or unanalyzable, we process the form as a whole.

The English past tense studies and the studies of left-hemisphere impaired patients give us an idea of how the lexicon is organized. As Marslen-Wilson and Tyler

(2007) said "identifying inflectional morphemes is a major priority for the language processing system" (p.831). Regular past tense verbs are decomposed into the word stem and the past tense morphemes when we process the words, whereas irregular past tense verbs are processed as a whole. Even though the suffix has various pronunciations, we decompose to a stem and a suffix.

To summarize this section, the differences in underlying representations are reflected in behavioral data. Although the evidence is indirect, it seems that people treat the suffix alternants, which are argued to have a single underlying representation or not, differently. Acquisition of the Korean nominative case marker in §2.4.1 shows that children treat the alternants differently. The nominative case marker is argued to favor the MAT-style analysis, yet the markers are not acquired at the same time and the errors that the children make conform to one of the phonologically-conditioned rules. §2.4.2 shows the acquisition of the English plural suffix, which favors the GP-style analysis. It seems that the English-speaking children learned the phonological transformation rule and the devoicing rule, and were able to apply the rules across the morphemes and also to new words. Remember that the children did not seem to learn all of the phonological transformation rules for the suffix at the same time. This can be argued to support the GP-style analysis in that they posited a single underlying representation and they were in the process of learning the phonological transformation rules. §2.4.3 shows that regular and irregular past tense verbs are processed differently. Regular past tense verbs go through morpho-phonological analysis, but irregular past tense verbs are processed as a whole because they are phonologically unanalyzable. As a result, we can infer that regular verbs are organized with a stem word and a past tense morpheme with phonological transformation rules, and irregular verbs are listed with the past tense form. Based on this, we can further infer that when the suffix or the alternant is phonologically unanalyzable, it is possible that we might process the alternant as a whole. In other words, it is possible that we list the alternants as their own form in the lexicon when they are phonologically unanalyzable.

The studies that I reviewed here do not directly ask whether there are any

differences in acquisition or processing when there is arguably a single underlying representation. Thus, I used the artificial grammar learning paradigm to study the question. The next section reviews the artificial grammar learning paradigm for learning phonological patterns.

2.5 Artificial Grammar Learning (AGL) Paradigm

The above mentioned questions are studied by using the artificial grammar learning (AGL) paradigm. The AGL paradigm consists of an exposure phase and a test phase. In the exposure phase, participants are exposed to grammars of interest in a limited period of time. In the test phase, participants are tested on how well they learned the patterns in the exposure phase and usually there is no negative feedback. Researchers often compare two grammars or patterns which minimally differ from each other. The important thing to point out about the AGL paradigm is that these studies are comparative. If participants tend to learn one pattern better than another, then the study is informative. AGL studies give us information about what we learn better.

The assumption behind the AGL paradigm is that the learning mechanism investigated in the AGL paradigm and natural language are shared (Folia et al., 2010; Gómez and Gerken, 2000; Petersson et al., 2004; Reber, 1967; Silva et al., 2017). However, the AGL paradigm is often criticized in that it is very different from learning a natural language. Children are exposed to multiple other aspects of language (e.g., syntax, phonology, and semantics) and social interaction than when learning a single aspect of a language in an experimental setting. Moreover, children are exposed to a language over a very long period of time. In other words, they are not exposed to a list of words repeatedly within a short period of time as in the exposure phase of the AGL paradigm.

Despite the differences, the advantage of the paradigm is that it can easily control the complexity of natural language acquisition. The complex process of acquisition makes it more difficult to investigate the mechanism of learning a language. The AGL

paradigm overcomes the difficulty by controlling various linguistic elements, including prior knowledge of a language. Thus, the AGL paradigm allows us to study each linguistic aspect seperately. There are many studies on the learnability of grammar (e.g., Marcus, 1999; Öttl et al., 2015) and on phonology (e.g., Finley, 2017; Lai, 2015; Peperkamp et al., 2006) that use the AGL paradigm.

In addition, the paradigm can be used for both adults (e.g., Finley and Badecker, 2010; Finley, 2015; Lai, 2012, 2015; Pycha et al., 2003) and infants (Chambers et al., 2003; Marcus, 1999; Seidl and Buckley, 2005; Seidl et al., 2009; White and Sundara, 2014). With brief exposure of patterns, both adults and infants show the same results in the learning (e.g., Chambers et al., 2003; Onishi, 2002). Therefore, even though the learning in the experimental setting is not exactly same as the natural language acquisition, the AGL paradigm has been argued to shed light on the various aspects of natural language acquisition.

2.5.1 AGL in Phonotactics

With brief exposure to new sound patterns, infants and adults are able to learn the grammar and the phonotactics of an artificial language (Chambers et al., 2003; Marcus, 1999; Onishi, 2002; Seidl et al., 2009). Learning the grammar or the phonotactics of a sound pattern shows the ability to abstract and generalize. In Marcus (1999), 7-month-old infants were able to learn and then generalize grammars such as ABA or ABB to new sounds or syllables. The ABA refers to the grammar where the first syllable and the last syllable are the same and the ABB refers to the grammar where the last two syllables are the same in three syllable words. For example, infants in the ABA condition were exposed to "ga ti ga", and were tested on new words such as "wo fe wo'," the legal pattern, or "wo fe fe," the illegal pattern. Infants looked longer at the illegal pattern, which means that they were able to generalize the grammar in the training phase and extend it to the new sounds or syllables. In Chambers et al. (2003), 16.5-month-old infants were able to learn phonotactic regularities that were not in their native language (English), and to generalize those regularities to new

35

syllables. For example, the infants were trained with /bæp/ as a legal sound pattern and /pæb/ as an illegal pattern. When they were tested on new legal patterns, the infants listened longer to the illegal patterns than to new legal patterns, which means that they noticed the difference and indicates that they learned the phonotactics of the artificial language and were able to generalize. Exposure to the training items showed the same results with adults (Onishi, 2002).

Learning of phonotactics is similar to learning the allophones of a language. Seidl et al. (2009) studied 11-month-old French-learning infants and 4- and 11-month-old English-learning infants. Nasal vowels are phonemic in French, but they are allophonic in English. Using oral vowels and nasal vowels, Seidl et al. trained 11-month-old French-learning and English-learning infants on phonotactics. The training items were CVC syllables, in which the vowel is either oral or nasal. When the oral vowel occurs, it is followed by stop consonants, and when the nasal vowel occurs, it is followed by fricatives. Test items included both legal and illegal (e.g., oral–fricative) sequences. French-learning infants were able to learn the phonotactic pattern and generalize, but English-learning infants did not because the nasal vowel is phonemic in French but not in English. However, 4-month-old English-learning infants were able to learn the phonotactic restrictions and generalize. The results indicate that 4-month-olds can generalize to the sound pattern and that the acquisition of allophones occurs between 4 and 11 months of age.

2.5.2 AGL in Phonological and Phonetic naturalness

In line with the studies on the learnability of phonotactics, there are ample studies on the role of phonological or phonetic naturalness in learning. Phonological or phonetic naturalness can be measured by phonological features involved or phonetic plausibility (Moreton and Pater, 2012b; Wilson, 2006). One hypothesis behind the studies is that people are biased to learn sound patterns that are more phonologically or phonetically natural, which is known as substantive bias (Wilson, 2006). However, results from various studies show mixed results. As reviewed in Moreton and

Pater (2012b), the learning of phonetic substance is "elusive and unreliable" (p.702) or "weaker" (p.709) than the learning of formal structure or complexity. There seems to be a preference for phonologically or phonetically natural patterns (Pycha et al., 2003; Finley, 2012, 2017; Wilson, 2003, 2006; White and Sundara, 2014), but people can also learn less natural patterns (Seidl and Buckley, 2005; Peperkamp and Dupoux, 2007; Linzen and Gallagher, 2014).

Studies that support the substantive bias hypothesis focus on assimilation/dissimilation and harmony/disharmony. Finley (2012) showed that people were better at learning phonetically motivated vowel harmony patterns. The study compared round vowel harmony in mid vowels, [e, o], and high vowels, [i, u]. Since mid vowels are perceptually weak or robust (Kaun, 2004), the mid vowels serve as triggers for harmony. Participants who were trained with the mid vowels were able to generalize the round harmony to high vowels, but participants who were trained with the high vowels were not able to generalize the round harmony to mid vowels.

In Wilson (2003), participants were better at learning assimilation and dissimilation processes than random processes. In that study, the suffixes alternated between [-la] and [-na]. In the assimilation condition, the suffix [-na] was selected when the last consonant in the stem was nasal; otherwise [-la] was selected. In the dissimilation condition, the suffix [-la] was selected when the last consonant in the stem was nasal; otherwise, [-na] was selected. The assimilation condition was compared to the first random condition, where the suffix [-na] was selected when the last consonant in the stem was dorsal sounds; otherwise [-la] was selected. The dissimilation condition was compared to the second random condition where the suffix [-la] was selected when the last consonant in the stem was dorsal sounds; otherwise [-na] was selected. Participants were better at learning the assimilation and dissimilation condition than the random conditions.

Participants in Pycha et al. (2003) were also better at learning the harmony and disharmony process than the arbitrary process. Both the harmony and disharmony conditions involved a feature, [back], for harmony and disharmony. However, in

the arbitrary condition, the feature [lax] also played a role along with [back]. Participants learned both of the harmony and disharmony conditions but not the arbitrary condition. However, Wilson (2003) and Pycha et al. (2003) did not show any learning difference between assimilation/harmony and dissimilation/disharmony.

Metathesis based on syllable structure also supports the substantive bias hypothesis (Finley, 2017). When two syllables are pronounced together, there are two options: one can pronounce the syllables by just adding them, or one can pronounce with metathesis (e.g., /ar/ + /ge/ → [arge] or [agre]). Since English syllabification allows and prefers complex onsets, one should prefer the metathesis version [agre] because it can be syllabified to [a.gre]. The form [arge] should not be preferred because *[a.rge] is not a possible onset in English, and [ar.ge] is not structurally preferred in terms of the Maximal Onset Principle (Selkirk, 1982). Participants were better at learning the metathesis pattern when the metathesis was structurally and perceptually motivated.

However, people do not learn only phonologically or phonetically natural patterns. In Seidl and Buckley (2005), infants learned phonetically natural and unnatural phonotactic patterns equally well. Infants between 8.5- and 9.5-months-old were exposed to a phonetically natural pattern and an arbitrary pattern. In the phonetically natural pattern, fricatives and affricates occur only intervocalically (intervocalic spirantization) or are followed by round vowels, and coronal consonants are followed by front vowels (consonant-vowel assimilation). In the arbitrary condition, fricatives and affricates occur word initially, labial consonants are followed by high vowels, and coronal consonants are followed by mid vowels; these are unattested patterns. There was no preference between the phonetically natural and unnatural phonotactic patterns, which indicates that they were able to learn the two patterns equally well.

In Peperkamp and Dupoux (2007), adults learned a phonetically unnatural group of allophones. For example, participants were trained on the alternations of [p]∼[ʒ], and [g]∼[f] intervocalically. In the test session, they used a picture-matching forced-choice task, and the participants had to infer that /z/ alternates with [t] or

38

/d/ alternates with [s] . These allophone classes do not form natural classes, yet participants were able to map the allophones successfully. When they changed the testing method to a picture-naming task in Peperkamp et al. (2006), and the participants had to pronounce the allophones, they performed better with the natural group of allophones. Linzen and Gallagher (2014) also showed that unattested patterns or phonetically unnatural patterns can be learned with enough data or exposure to the items.

The mixed results of these studies show that there seems to be a preference for learning phonologically natural patterns, but it is not impossible to learn phonetically unnatural patterns in experimental settings.

2.5.3   AGL in Complexity

Differences in learning may also come from differences in formal or computational complexity. People learn less complex patterns better than more complex patterns in artificial phonological learning (Moreton and Pater, 2012a). Complexity can be measured in various ways: number of phonological features involved, relationship between the features, and with logic (Subregular Hierarchy (McNaughton and Papert, 1971; Rogers and Pullum, 2011; Rogers et al., 2013)).

Some of the studies mentioned above also show support for differences in learning based on the complexity of the features. Participants in Pycha et al. (2003), for example, learned the harmony and disharmony pattern, which changed only one feature [back], better than the arbitrary pattern, which changed two features, [back] and [lax]. Finley (2015) studied nonadjacent dependencies with vowel harmony. Vowel harmony occurs in adjacent vowels, but when there is a transparent vowel, the vowel can be skipped, creating a second-order long-distance dependency. Participants in the study were better at learning the vowel harmony pattern without transparent vowels. But this did not mean that the transparent vowel harmony patterns were unlearnable. People learned the transparent vowel harmony pattern with additional exposure.

Not only the number or relationships of the features, but also the computational complexity, plays a role. Finley and Badecker (2010) showed that participants were better at learning an attested vowel harmony pattern than an unattested pattern, which is also computationally more complex (see Gainor et al., 2012, for more details). The attested patterns were vowel harmony with directions either from left to right or from right to left. The unattested pattern was the 'majority rules' pattern, where the spreading feature value is determined by whether it is a majority in the underlying representation. For example, imagine a language with a vowel harmony pattern where the vowels agree in the feature [back]. If a word has $/+ - -/$ feature of [back], the [–back] spreads from right to left as $[- - -]$. But if a word was $/- - +/$, the feature spreads from left to right. In short, both patterns change one feature with a different direction. Subjects were able to learn the directional vowel harmony pattern, but not the majority rules pattern.

Lai (2012, 2015) showed the learnability of phonotactics based on consonant harmony patterns. In these studies, complexity was measured in terms of the Subregular Hierarchy. The attested consonant harmony pattern was computationally less complex than the unattested consonant harmony pattern, and participants were better at learning the attested consonant harmony pattern. The attested pattern was a sibilant harmony pattern where all the sibilants had to agree in anteriority in a word, regardless of the distance between the two segments. The pattern will allow [s ... s ... s] and [ʃ ... ʃ ... ʃ] patterns, but not [s ... ʃ ... s] or [ʃ ... s ... ʃ]. The unattested pattern was called First-Last harmony where only the first and the last segment had to agree in their anteriority. The pattern will allow [s ... ʃ ... s] and [ʃ ... s ... ʃ] because the anteriority of the first and last segment agree in both words, but not [s ... ʃ ... ʃ] or [ʃ ... s ... s]. It is not always the case that computationally less complex patterns are attested patterns.

In Hwangbo (2015), the unattested vowel harmony pattern was less complex than the attested vowel harmony pattern, and participants learned the less complex vowel harmony pattern better. In the study, the suffix was selected according to a

vowel in the stem word. In the unattested pattern, the suffix with a front vowel was selected when there was at least one front vowel in the stem, while in the attested pattern, the suffix with a front vowel was selected when the last vowel in the stem was a front vowel. Regardless of attestedness, participants were better at learning the less complex pattern consistently.

In this section, I reviewed various studies that implemented the artificial grammar learning paradigm. Overall, researchers have used the AGL paradigm to study various aspects of learning sound patterns, such as phonotactics, allophones, the relationship with phonologically or phonetically natural sound patterns, and formal complexity. These various aspects reflect the learning of natural language in some sense, and thus I also use this paradigm for this dissertation.

2.6 Chapter Summary

To summarize this chapter, I introduced GP and MAT in broad terms in §2.1. The main difference between GP and MAT is the existence of phonology. In GP, there is a single underlying representation for the alternants, and the surface forms result from phonological transformation, which constitutes the phonological module. In MAT, however, the alternants are all listed in the lexicon and the morphological selection rules select the appropriate form; there is no phonological transformation. I reviewed language data supporting these theories in §2.2. There is ample data to support GP, and some data to support MAT. I also reviewed behavioral studies in §2.4. Although the behavioral studies do not directly reflect either theory, it seems there are some differences in treating the alternants. Lastly, in §2.5, I reviewed AGL studies showing that the paradigm captures some aspects of the learning of natural languages.

Although there are ample studies, the behavioral studies and the AGL studies do not directly inquire into the relationship between the underlying representations and the learning. Therefore, in the next chapter I pose a research question regarding the relationship between underlying representations and their learnability. To be more

specific, the underlying representations of morpheme alternants are in question. Will there be any difference in learning two morpho-phonological patterns, with one more amenable to GP analysis like the Kinyarwanda verb root forms and one more amenable to a MAT analysis like the Korean nominative? This question is based on the studies reviewed in this chapter. Behavioral studies show that people treat the morphemes differently. If this is the case, the difference should be reflected in the right experimental study as well. Although the previous behavioral and experimental studies tapped into the question of learnability in terms of phonology, the fundamental question of underlying representations and phonology, and their learnability, has not yet been studied directly. Because the AGL paradigm reflects aspects of phonological learning, I use it in this dissertation.

Chapter 3

THE RESEARCH QUESTION AND THE ARTIFICIAL GRAMMAR

In this chapter, I propose three models of learning alongside with the research question and hypotheses in §3.1 and present the artificial grammars for this dissertation in §3.2 and §3.3. Lastly, I discuss the problem of phonotactics in §3.4 followed by the chapter summary in §3.5.

3.1 The Research Question and Hypotheses

The question arises from the difference among what the models predict when learning a new morpho-phonological pattern. When there are two morpho-phonological patterns, which differ in the likelihood of phonological processes, will there be any difference in learning? To be more specific, one morpho-phonological pattern is more likely to be analyzed with phonological processes when positing underlying representations, and another morpho-phonological pattern is less likely to analyzed with phonological processes.

When learning a morpho-phonological pattern, there are at least three types of models under GP. Either phonological learning or morphological learning is activated before the other, or they are both activated simultaneously. Each model is outlined in (3.1).

(3.1) a. Phonology before morphology (P-FIRST) model :
         Phonological learning is activated before morphological learning.

       b. Morphology before phonology (M-FIRST) model :
         Morphological learning is activated before phonological learning.

c. Simultaneous (SIMULTANEOUS) model :

Phonological learning and morphological learning are activated simultaneously.

The Phonology before morphology (P-FIRST) model activates phonological learning before morphological learning. The phonological module is considered to be part of human language faculty (e.g., Berent, 2013; Fitch et al., 2005; Hauser et al., 2002), and thus, humans instinctively posit a phonological analysis whenever possible. Thus, the patterns will be easier to learn if the alternants are phonologically motivated. Based on the phonological analysis, the model posits the underlying representation as a single form or multiple forms. If the alternants can be abstracted into a single underlying representation based on phonological analyses, then the model posits a single underlying representation and phonological transformations. If not, the model posits multiple underlying representations with phonologically-conditioned rules.

The Morphology before phonology (M-FIRST) model activates morphological learning before phonological learning. Thus, the M-FIRST model first lists the surface representations as the underlying representations before phonological analyses, and then the model will categorize the words into groups with morphological selection rules. Once the morphological grammar is obtained, a phonological grammar can be inferred. If it is possible to abstract the phonology of the allomorphs then the model posits a single underlying representation. If not, the model keeps the lists of morphemes with morphological selection rules.

In the Simultaneous (SIMULTANEOUS) model, phonology and morphology are activated simultaneously and compete with each other. Simultaneous or competing grammars are not a novel concept (e.g., Yang, 2000), in that one rule applies as a default and the other rule competes with the default rule. Adopting the concept, I assume either phonology and morphology applies and the other competes with the other, and the module which has a better solution wins. There are many ways to specify the competition, for example, non-linguistic cognitive learning can also be considered

as one of the modules to compete. However, it is unclear what to specify as competing modules and which competes against each other. In this dissertation, therefore, this model is not going to be discussed in detail.

To summarize, the theories and models we have discussed, (3.2) shows the relationships between the two theories and models. Broadly, there are two views: MAT and GP. Under GP, there are three models. The three models differ in terms of which module is activated before the other, or activated simultaneously.

(3.2)   a. MAT

      b. GP

         i. P-First model

        ii. M-First model

       iii. Simultaneous model

Turning from the models to the research question, each model predicts different outcomes in learning. The question, whether the two morpho-phonological patterns, which differ in the phonological processes, will show any difference in learning, is examined with tje AGL paradigm in this dissertation.

In the AGL paradigm, I set up two grammars that I call the GP grammar and the MAT grammar. The details of each grammar are discussed in §3.2 and §3.3. Briefly, the morphemes alternate based on the phonological process in the GP grammar, and thus the alternants are more likely to be analyzed as a single underlying representation. In the MAT grammar, there is no phonological motivation for the morpheme alternants, which is more like phonologically-conditioned allomorphy. Therefore, in the MAT grammar, the alternants are more likely to be listed as their own underlying representations. The hypothesis following this logic is stated in (3.3).

(3.3)   a. Null Hypothesis:

        There will be no difference in learning the two grammars.

      b. Experimental Hypothesis:

        There will be a difference in learning the two grammars.

The null hypothesis (3.3a) is predicted by MAT because it is just listing all the allomorphs as the underlying representation and categorizing which suffix goes with which category of words. MAT does not predict any difference in learning because no phonological analysis is required. Assuming an equal amount of computation in each type of learning (phonological and morphological) and in each grammar, this hypothesis is also predicted by the M-First model because both grammars require morphological learning and phonological learning.

The experimental hypothesis (3.3b) is predicted by the P-First model. To be more specific, the P-First model predicts that when the phonological learning abstracts and posits a single underlying representation, morphological learning is not required. In other words, it should be easier to analyze and learn phonological processes if the alternations are phonologically motivated than the phonologically unmotivated alternations. Therefore, the P-First model predicts that there will be a difference in learning the patterns, specifically, in favor of learning the GP grammar over the MAT grammar.

The research question asks directly whether the representational differences in allomorphs result in differences in learning. The question is examined with the AGL paradigm, and in the next two sections I present the GP grammar in §3.2 and the MAT grammar in §3.3.

3.2   The GP Grammar

As briefly mentioned in the previous section, the alternants in the GP grammar are based on the phonological process, and thus, the alternants are more likely to be analyzed as a single underlying representation.

The GP grammar is based on the Danish vowel lowering process (Basbøll, 2005; Trubetzkoy, 1969). In Danish, vowels lower before or after a rhotic sound; this is also known as r-colored vowels. The vowel inventory of Danish vowels outside of the /r/-context is shown in Table 3.1.

	Front		Back
	unrounded	rounded	
	i	y	u
	e	ø	o
	ɛ	œ	ɔ
	a		
	ɑ		ɒ

Table 3.1: Danish vowel phonemes outside of the /r/-context (from Basbøll (2005))

The vowel /e/ is realized as [ɛ] after /r/, /ɛ/ is realized as [æ] before /r/ and as [a] after /r/, and /ø/ is realized as [œ] before and after /r/ (Basbøll, 2005). For example, the word *bær* 'berry' is pronounced as [ˈb̥æɐ̯] and the word *bræk* 'vomiting' is pronounced as [ˈb̥ʁaɡ̊].

For simplicity, in this experiment the vowel lowering process takes place only after /r/ and on the front vowels. Based on the traditional rule-writing system, the lowering process of the front vowels for this dissertation is represented as (3.4).

(3.4)   Vowel lowering: Front vowels lower when it follows /r/.

a. $[-high, -low, +front] \rightarrow [+low]/$ r ____

b. $[+high, -low, +front] \rightarrow [-high] /$ r ____

The GP grammar follows the vowel lowering process in (3.4). The vowels in the GP grammar are [i, e, æ, u, o], and the vowel lowering applies only on the front vowels, [i, e, æ]. Therefore, after /r/ the underlying high front vowel /i/ surfaces as the mid front vowel [e], and the underlying mid front vowel /e/ surfaces as the low front vowel [æ], as shown in (3.5). Note that the high front vowel [i] does not lower to [æ].

(3.5)   a. /e/ → [æ]/r____

         b. /i/ → [e]/r____

Participants were not exposed to the sequence [ri] because the vowel was always lowered on the surface form in the experiment. In other words, participants were exposed to sequences like [re] and [ræ], but not [ri]. Examples of the processes are shown in (3.6).

(3.6)   a. /ri/ → [re]

         b. /re/ → [ræ]

There were two pairs of suffixes: the [-et]~[-æt] pair and the [-em]~[-æm] pair. Following GP, underlying representations of both allomorphs ought to be analyzed as /-et/ and /-em/. In other words, the GP-grammar shows the vowel lowering process discussed above in the word stems as well as in the suffixes. Examples are given in (3.7), and the full lists of items are in Appendix A and Appendix B.

(3.7)   a. /bos-et/ → [bos-et]

             /bos-em/ → [bos-em]

         b. /gur-et/ → [gur-æt]

             /gur-em/ → [gur-æm]

         c. /res-et/ → [ræs-et]

             /res-em/ → [ræs-em]

         d. /rir-et/→ [rer-æt]

             /rir-em/→ [rer-æm]

Words that do not contain /r/, such as (3.7a), do not show any phonological process, whereas words that contain /r/ show the phonological process in the stem or in the suffix (or both), as in (3.7b), (3.7c), and (3.7d). Words such as (3.7b) end with /r/ and thus the suffix alternates. Words such as (3.7c) begin with /r/ and thus the following vowel lowers. Words such as (3.7d) begin and end with /r/ and thus the vowels following /r/ lower in both the stem and the suffix.

The four types of stems were CVC, CVr, rVC, and rVr, where C stands for consonants other than /r/, and V stands for the vowels. (3.8) shows the allomorphs that the stem occurs with. The details of the experiments are discussed in Chapter 4 and Chapter 5.

(3.8)   a. CVC-et/em

            b. CVr-æt/æm

            c. rVC-et/em

            d. rVr-æt/æm

In the GP grammar there is more likely to be a single underlying representation for the allomorphs because the artificial grammar consistently shows the phonological transformation. Therefore, if phonological learning is activated before morphological learning as the P-First model predicts, morphological learning is not required to be activated, and thus, participants will be better at learning the patterns. The M-First model, however, predicts that there will be no difference in learning the two grammars. If people's morphological learning is activated before phonology, one lists the morphemes and categorizes the words and figures out which words go with which suffix as 'when the word ends with [r], the lower suffix [-æt] or [-æm] is selected, and otherwise [-et] or [-em]'. This categorizes the stem types into two: one that ends with /r/ and another that ends with other consonants. Then, the model proceeds to phonological learning and abstracts the phonology of the pattern. Therefore, the M-First model requires both types of learning, which predicts differently from the P-First model.

The MAT-style analysis, which does not include any phonological analysis, however, predicts that it is just categorizing words into two groups and figuring out the phonologically-conditioned selection rule.

Table 3.2 summarizes the predictions of each theory and model. The morphological learning in the P-First model is left empty because the analysis is done in the phonological learning. Each learning assumes an equal amount of computation, thus,

the P-First model ends the learning earlier than the M-First model. As a result, the P-First model predicts that people will learn the GP grammar better.

Table 3.2: Learning prediction of the GP grammar.

MAT		Morpheme listing and categorization
GP P-First		
	Phonological learning:	Abstracts phonology with a single UR
	Morphological learning:	—
M-First		
	Morphological learning:	Morpheme listing and categorization
	Phonological learning:	Abstracts phonology with a single UR

The [-et]∼[-æt] pair was presented in Experiment 1A, and the [-em]∼[-æm] pair was presented in Experiment 1B. In these experiments, the first pair,[-et]∼[-æt], was introduced as a plural suffix and the second pair, [-em]∼[-æm], was introduced as a possessive suffix. In Experiment 2, both pairs were presented to the participants.

3.3 The MAT Grammar

Matching the GP grammar, the MAT grammar shows phonologically-conditioned allomorphy for words ending in /r/. The MAT grammar is less likely to be analyzable with phonological transformations. It is based on a phonologically-conditioned allomorphy (e.g., the Korean nominative case marker described in §2.2.4). Thus, both alternants are more likely to have underlying forms of their own.

The phonologically-conditioned selection rule is similar to the GP grammar. There are two pairs of suffix alternants: [-ut]∼[-æt] and [-up]∼[-un]. The suffix alternates when a word ends with /r/. The rule is stated in (3.9).

(3.9)   Plural suffix selection rules:

    a. When a word ends with /r/, /-æt/ is selected as a suffix.

    b. When a word ends with a consonant other than /r/, /-ut/ is selected as a suffix.

(3.10)   Possessive suffix selection rules:

    a. When a word ends with /r/, /-un/ is selected as a suffix.

    b. When a word ends with a consonant other than /r/, /-up/ is selected as a suffix.

Note that there is no phonological transformation in this grammar. Therefore, all vowels occur after /r/, and thus, participants are exposed to the [ri] sequence, unlike in the GP grammar. In the MAT grammar, the underlying representations of all of the stem words and suffixes are the same as the surface forms.

Unlike the GP grammar, the alternants [-ut]∼[-æt] and [-up]∼[-un] are less likely to be analyzed into a single underlying representation. In other words, the alternations between [-ut]∼[-æt] and [-up]∼[-un] have no phonological motivation. If one tries to analyze the MAT grammar in the GP-style, one will run into a problem: the MAT grammar has more features to change, and, the changes have no phonological motivation. For example, for the [-ut]∼[-æt] pair, one has to change [+back, +high, −low] to [−back, −high, +low] when the word ends with /r/. For the [-up]∼[-un] pair, one has to change place, voicing, and nasality features when the word ends with /r/. Therefore, the MAT grammar is more likely to be analyzed with all the alternants listed in the lexicon with the selection rules in (3.9) and (3.10). Examples are given in (3.11).

(3.11)   a. /bos-ut/ → [bos-ut]

         /bos-up/ → [bos-up]

    b. /gur-æt/ → [gur-æt]

         /gur-un/ → [gur-un]

c.　/res-ut/ → [res-ut]

　　　/res-up/ → [res-up]

　d.　/rir-æt/→ [rir-æt]

　　　/rir-un/→ [rir-un]

According to the selection rules in (3.9) and (3.10), words ending in /r/ select the suffix /-æt/ or /-un/, as shown in (3.11b) and (3.11d), otherwise they select /-ut/ or /-up/, as shown in (3.11a) and (3.11d). All the underlying representations are the same as the surface forms.

　　　As in the GP grammar there were four stem types. (3.12) shows the allomorphs that the stem occurs with.

(3.12)　a.　CVC -ut/up

　　　　b.　CVr - æt/un

　　　　c.　rVC - ut/up

　　　　d.　rVr - æt/un

　　　As mentioned above, the MAT grammar is more likely to have multiple underlying representations such as /-ut/, /-up/, /-æt/, and /-un/. The MAT-style analysis predicts that it is categorizing words into two groups, the ones that end with /r/ and others ending with other consonants, and identifying out the phonologically-conditioned selection rules. Therefore, there will be no difference in learning the GP grammar and the MAT grammar.

　　　The GP-style analysis predicts a difficulty in learning the MAT grammar because the models require both phonological learning and morphological learning. The P-First model predicts that people will have more difficulty learning the MAT grammar because the alternations are not phonologically motivated, and thus, the model has to activate the morphological learning unlike the GP grammar. Phonological learning will find that the patterns cannot be generalized with phonology, then, the model moves on to morphological learning and the alternants are listed as their own representations.

The M-First model also predicts a difference. The model first lists and categorizes the alternants, and then, activates phonological learning. Phonological learning finds out that the alternants cannot be abstracted, and thus, the morphemes are listed in their own representations. Both models activate both phonological learning and morphological learning.

Table 3.3 summarizes the learning prediction of the MAT grammar of each theories. Phonological learning in the P-First model and the M-First model is marked with an X-mark (✗) because the module cannot abstract phonology from the patterns. Again, each type of learning assumes an equal amount of computation in each model.

Table 3.3: Learning prediction of the MAT grammar.

MAT	Morpheme listing and categorization
GP P-First	
Phonological learning:	✗
Morphological learning:	Morpheme listing and categorization
M-First	
Morphological learning:	Morpheme listing and categorization
Phonological learning:	✗

The [-ut]∼[-æt] pair was presented in Experiment 1A, and the [-up]∼[-un] pair was presented in Experiment 1B. The suffixes were also described with two different meanings: the [-ut]∼[-æt] pair as a plural suffix and the [-up]∼[-un] pair as a possessive suffix. Both pairs were presented in Experiment 2.

3.4 Why a Vowel Lowering Process?

In this section, I discuss why I adopted the vowel lowering process for this dissertation.

The vowel lowering process was adopted to avoid possible confounding due to phonotactics. As reviewed in §2.5.1, people learn phonotactics with brief exposure in the AGL paradigm. The purpose of this dissertation is to examine the role of phonology when learning a new morpho-phonological pattern, and not phonotactics. Thus, the patterns should not be learnable by phonotactics.

Based on the nature of the artificial grammar learning paradigm with a two-alternative forced-choice task, participants are exposed to both well-formed and ill-formed choices at the same time. Therefore, it is possible to use phonotactic information to select the correct word forms. If this were the case, no inference about lexical storage and transformations could be drawn.

Here is an example of experimental stimuli that uses phonotactics rather than learning the lexical representation. Let's suppose there is an allomorph that alternates between V and CV, where V stands for vowels and C stands for consonants. Like the Korean nominative case marker, we can posit an alternation between [-i]∼[-ki] for the GP grammar and [-a]∼[-ki] for the MAT grammar. In the GP grammar, one can posit a single underlying representation /ki/, where the consonant /k/ deletes after another consonant to avoid consonant cluster; thus, the words show CVC-*i* and CV-*ki*. In the MAT grammar, the suffix may alternate between [-a]∼[-ki], where the [-a] form is selected when the word ends with a consonant, and otherwise [-ki] is selected. This is a phonologically-conditioned allomorphy, where the phonological transformation between the allomorphs has no phonological motivation. Therefore, the legal forms are CVC-*i* and CV-*ki* in the GP grammar, and CVC-*a* and CV-*ki* in the MAT grammar. In short, participants in both grammars are trained on CVC-V and CV-CV syllables.

In the test phase, however, participants must choose between the legal form and the illegal form in a two-alternative forced-choice task. That is, participants have to choose the correct form between CVC-V and *CVC-CV or CV-CV and *CV-V. For example, in the GP condition, participants must select the correct answer between CVC-*i* and *CVC-*ki*. This is where the phonotactics problem is found. Participants never heard *CVC-CV or *CV-V forms in the training session. Therefore, it is possible

for participants to choose the correct answer, CVC-*i* or CV-*ki* just because the word form was illegal in the other option, and not because they learned the lexical representation or phonological transformation. Therefore, in this dissertation, I do not use such patterns, but use the vowel lowering process to eliminate the possibility of using phonotactics.

As mentioned in the previous sections, I adopt the vowel lowering process from Danish. The front vowels lower after /r/ in the GP grammar, which indicates that participants hear all the front vowels after [r] except the high front vowel [i]. In other words, participants hear both [re] and [ræ] sequences in the stem such as [rep] and [ræp] or [rer] and [rær].

The vowel lowering process does not raise the phonotactics problem in the corresponding suffix alternations between [-et]∼[-æt] and [-em]∼[æm]. Participants are trained on CVC-[et] and CVr-[æt], for example. Participants are exposed to the [ræ] sequence in the stems, such as [ræp], and thus the suffixed form does not raise any phonotactics problem.

In the test phase, therefore, the two-alternatives do not violate phonotactics. Example (3.13) shows both legal forms and illegal forms that participants hear during the test phase.

(3.13)   Legal forms   Illegal forms
   a.   rep-et   rep-æt
   b.   ræp-et   ræp-æt
   c.   rer-æt   rer-et
   d.   rær-æt   rær-et

As example (3.13) shows, the two-alternatives in the forced-choice task do not present any phonotactics problem. By implementing the vowel lowering process, one cannot use any phonotactic information to select the correct answer, but must know the phonological process.

Based on the nature of a two-alternative forced-choice task, participants must be exposed to legal and illegal forms in the test phase. By eliminating the possibility of using phonotactics to select the answer, we can infer that participants are using their lexical memory or phonological knowledge to answer, rather than phonotactics which would possibly confound the purpose of this dissertation.

3.5 Chapter Summary

To summarize this chapter, I began by positing the research question in §3.1 and then introduced two artificial grammars in §3.2 and §3.3. Lastly, in §3.4 I discussed the problem of phonotactics, which justifies why I implemented a vowel lowering process in this dissertation.

The research question is based on the logic of possible models when learning morpho-phonological patterns. The first model predicts that people will activate phonological learning before morphological learning (the P-First model) and the second model predicts that people will activate morphological learning before phonological learning (the M-First model). Because phonological learning is activated before morphological learning in the P-First model, people will learn a morpho-phonological pattern easier if the alternation is phonologically motivated which requires only phonological learning. In the M-First model, both morphological and phonological learning is always required, thus, it predicts no difference in learning. A third model, Simultaneous, imagines a more complex interaction between phonology and morphology learning.

There were two artificial grammars, which differ only in the phonological process of vowel lowering, the GP grammar and the MAT grammar. The alternants in the GP grammar are based on the phonological process of vowel lowering, and thus more likely to be analyzed into a single underlying representation. The alternants in the MAT grammar are phonologically unmotivated, and thus more likely to be analyzed into multiple underlying representations with phonologically conditioned selection rules.

The hypotheses based on the models are as follows. MAT and the M-FIRST model predict that there will be no difference in learning the two grammars. MAT predicts no difference because it is categorizing words into groups and finding phonological conditions. The M-FIRST model predicts no difference because both grammars require both morphological learning and phonological learning.

The P-FIRST model predicts difference in learning. The model predicts that people will learn the GP grammar easily because the alternants are phonologically motivated, and thus, morphological learning is not required. It is assumed that an equal amount of computation is performed in phonological learning and morphological learning in each grammar.

Table 3.4 summarizes Table 3.2 and Table 3.3 in terms of what each theory and model is predicted to analyze for each grammar. The X-mark (✗) indicates that the grammar cannot be analyzed in the module, and the dash indicates that there is no further analysis in the module.

Table 3.4: Summary of how each model analyzes the GP grammar and the MAT grammar

		GP grammar	MAT grammar
MAT		Morpheme listing and categorization	Morpheme listing and categorization
GP	M-FIRST		
	Morphological learning:	Morpheme listing and categorization	Morpheme listing and categorization
	Phonological learning:	Abstracts phonology with a single UR	✗
	P-FIRST		
	Phonological learning:	Abstracts phonology with a single UR	✗
	Morphological learning:	–	Morpheme listing and categorization

Lastly, I discussed the problem of phonotactics. Due to the nature of two-alternative forced-choice tasks, participants are exposed to both legal and illegal forms in the test phase. To avoid any possibility of relying on phonotactics when selecting the answer, I use a vowel lowering process as the artificial grammar. During the training, participants are exposed to all of the sound sequences that appear in the illegal forms in the test session. Therefore, participants cannot use phonotactics to select an answer but must know lexical representations or phonological processes.

The next two chapters present three experiments, with their methods, results,

and discussions. The experiments are based on two grammars: the GP grammar and the MAT grammar. In Experiments 1A and 1B, participants learned one of the alternant pairs introduced in §3.2 and §3.3, and in Experiment 2, participants learned both pairs.

Chapter 4

EXPERIMENT 1

In this chapter, I present the methods–procedure and stimuli–, the results, and discussion of Experiment 1 which is composed of Experiment 1A and Experiment 1B. The experiment follows the grammar introduced in the previous chapter and examines whether there is any difference in learning the grammars when learning one pair of morpheme alternants. In Experiment 1A, participants in the GP condition learned [-et]∼[-æt], and participants in the MAT condition learned [-ut]∼[-æt]. In Experiment 1B, participants in the GP condition learned [-em]∼[-æm] and participants in the MAT condition learned [-up]∼[-un].

4.1 Experiment 1A: [-et]∼[-æt] and [-ut]∼[-æt]

4.1.1 Methods

4.1.1.1 Subject

A total of 40 subjects (20 in each condition) participated in the experiment (mean age: 19.7; 3 were males). Subjects were University of Delaware students who were native speakers of English and received a course credit as compensation.

4.1.1.2 Procedure

There were two conditions in the experiment: the GP condition and the MAT condition. Participants were randomly assigned to a condition. All stimuli were presented electronically using the E-Prime 2.0 software .(Psychology Software Tools, Inc., Pittsburgh, PA) on a Windows PC in a sound attenuating booth in the Phonetics and Phonology Lab at the University of Delaware.

Participants were instructed verbally about the general structure of the experiment and what the tasks were before the experiment. Participants were told that they are going to learn a new language that they never heard before, and they have to learn the words and the words with the plural suffix during the training session and then they will take a test following the training session.

During the experiment, written instructions were provided on the computer screen. The experiment was composed of two parts: the training session and the test session. In the training session, participants began with reading the instructions on the screen. They were instructed that they will learn a new language that they never heard before. To learn a language, they will hear a series of word pairs. The word pairs consisted of a stem word followed by the suffixed form. Participants were asked to listen to the word pair and then repeat the word and the suffixed form to learn the language. Participants listened to the word pairs while seeing '+' sign on the screen. After they heard the word pairs, they saw 'repeat the words you just heard.' The training was approximately 11 minutes. At the end of the training, participants saw 'this is the end of the training', and then they read instructions for the test session. Full instructions of the experiment screen are in Appendix C. Figure 4.1 and Figure 4.2 show screens that the participants saw during training session and the test session, respectively.
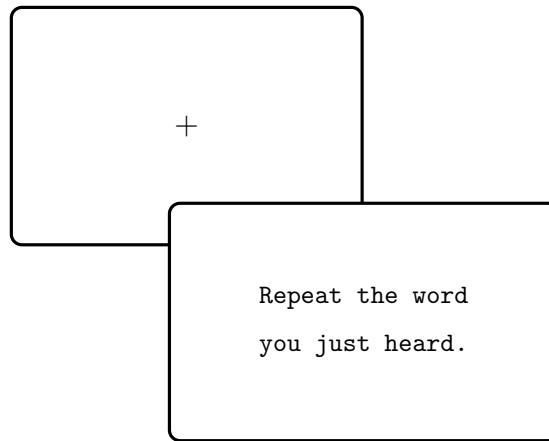
Figure 4.1: Example screens of training session
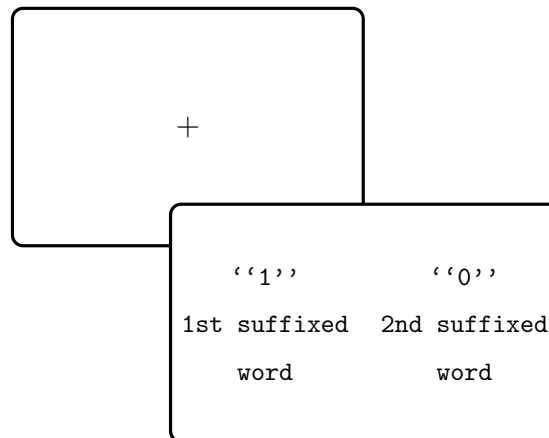
Figure 4.2: Example screens of test session

In the test session, participants began with reading instructions for the test. They were instructed that they will hear three words in a row: a stem word followed by two suffixed forms, and then choose the suffixed form that they think is grammatical or correct. Participants heard the words while seeing a '+' sign on the screen. Then,

the screen tells them to choose between the '1st suffixed word' and the '2nd suffixed word' by pressing the 1 or 0 key on the keyboard. One of the suffixed forms was correct according to the training condition. This was a two-alternative forced-choice task (2FA) where participants were asked to select the grammatical word form from two options, which were the two suffixed word forms.

The total duration of the experiment was approximately 30 minutes, and there was no break between the training and the test session.

4.1.1.3   Stimuli

All of the items included in the training and the test sessions were one syllable words. The consonants were [p, t, k, b, d, g, m, n, s, r], and the vowels were [i, e, æ, u, o]. There were four stem types, two of which ended with a rhotic and two of which ended with a consonant other than a rhotic. All four stem types were presented in the training session along with suffix corresponding to the condition. The suffixes presented with the stem types in each condition are presented in Table 4.1. The test session contained 'New items' and 'Old items'. New items were words that were not in the training session, whereas the Old Items were the same words from the training session. There was no new 'rVr' stem type due to the nature of the experiment, which had only the old items. The number of items is presented in Table 4.2.

The full list of training and test items are in Appendix A and Appendix B.

Table 4.1: Suffixes for the GP condition and the MAT condition in Experiment 1A.

Stem Types	Suffix	
	GP	MAT
CVC	-et	-ut
CVr	-æt	-æt
rVC	-et	-ut
rVr	-æt	-æt

Table 4.2: The number of training and test items in Experiment 1A.

| Stem Types | Training | | Test | | | |
			New Items		Old Items	
	GP	MAT	GP	MAT	GP	MAT
CVC	10	10	15	15	4	5
CVr	10	10	15	15	4	5
rVC	10	10	15	15	4	5
rVr	4	5	–	–	4	5

The suffixes were introduced as the plural suffix of the words. Each of the items in the training session were randomly repeated 5 times during the training session. The test items were also presented randomly. All items were recorded by a phonetically trained female native English speaker, and all the stimuli were provided in audio only.

4.1.2   Results

The results of New Items and Old Items were analyzed separately. The results show that participants were better at learning the CVr stem type in the New Items in the MAT condition but there was no difference in the Old Items.

The descriptive statistics given in Table 4.3 represent the proportion of the grammatical suffixes that were accurately selected. The item 'rep' was not included in the analysis due to experimenter's error.

Table 4.3: Descriptive statistics of the GP and MAT conditions in Experiment 1A. Rate of selecting the accurate suffix and standard error in Experiment 1A.

Stem Type	New		Old	
	GP	MAT	GP	MAT
CVC	0.837 (0.021)	0.873 (0.019)	0.900 (0.034)	0.870 (0.034)
CVr	0.577 (0.029)	0.743 (0.025)	0.613 (0.055)	0.800 (0.040)
rVC	0.832 (0.022)	0.857 (0.020)	0.883 (0.042)	0.913 (0.032)
rVr	–	–	0.638 (0.054)	0.770 (0.042)

Figure 4.3 shows the accuracy rate with the averages and standard errors for the New Items, and Figure 4.4 shows the average accuracy rate of each subject. The colored dot represents average of the accuracy rate with standard error.

Figure 4.3: Proportion of selecting the grammatical suffix for the New Items in Experiment 1A, with standard errors.

Figure 4.4: Proportion of selecting the grammatical suffix of each subject for New Items in Experiment 1A. Each dot represents for a subject. Colored dots represent average with standard errors.

Results were analyzed with a mixed-effect model with a binomial function using the *glmer* function included in the *lme4* package (Bates et al., 2015) in *R* (R Core Team, 2017).

The model included two random effects (subjects and test items) and two fixed effects (condition and stem types). There were two levels in the condition (GP and MAT), and there were four levels in the stem types (CVC, CVr, rVC, rVr). The subjects' responses were coded as '1' if they chose the grammatical suffix of the rhotic-ending words (i.e., the suffix [-æt] for the experiment) and '0' otherwise. In the analysis,

the CVC stem type in the GP condition was the reference level. Table 4.4 summarizes the fixed effects of the mixed-effects model for the New Items in the experiment. The effects include the between-subject variable (the MAT condition compared to the GP condition), and within-subject variables (CVr and rVC types compared to CVC stem type). The model also includes the interaction between the stem types and the conditions.

Table 4.4: Results of the fixed effects in the mixed-effects model for New Items in Experiment 1A.

	Estimate	Std. Error	z value	$\Pr(>\mid z \mid)$
(Intercept)	-1.864	0.252	-7.406	<0.001 ***
CVr	2.210	0.230	9.608	<0.001 ***
rVC	0.043	0.248	0.174	0.862
MAT	-0.279	0.350	-0.796	0.426
CVr×MAT	1.129	0.315	3.587	<0.001 ***
rVC×MAT	0.111	0.340	0.326	0.744

The baseline of the model (Intercept) is the log-odds of choosing the suffix related to the rhotic-ending words (i.e., [-æt]). The negative intercept indicates that the subjects are less likely to choose the suffix [-æt] in the CVC stem type in the GP condition. The log-odds of CVr indicate that the subjects are more likely to choose the suffix [-æt] in that stem type. The log-odds of rVC are not significantly different from the baseline. The log-odds of the condition (MAT) are not statistically significant ($p = 0.426$). The interaction of the stem type and the condition indicates that subjects choose the conforming suffix [-æt] more often in the MAT condition with the CVr stem type ($p = 0.000$) but not with the rVC stem type ($p = 0.744$), meaning that the

difference between the GP condition and the MAT condition in the CVr stem type is statistically significant but not in other stem types.

A one-sample t-test was conducted to determine whether the accuracy rate of the CVr stem type in the GP condition for the New Items ($M = 0.577$, $SD = 0.314$) was statistically different from a chance level, $t(19) = -704.7$, $p < 0.001$.

Figure 4.5 shows the proportion of selecting the conforming suffixes for the Old Items in Experiment 1A and Figure 4.6 shows the average accuracy rate of each subject. The colored dot represents the average accuracy rate in each stem type with standard errors.
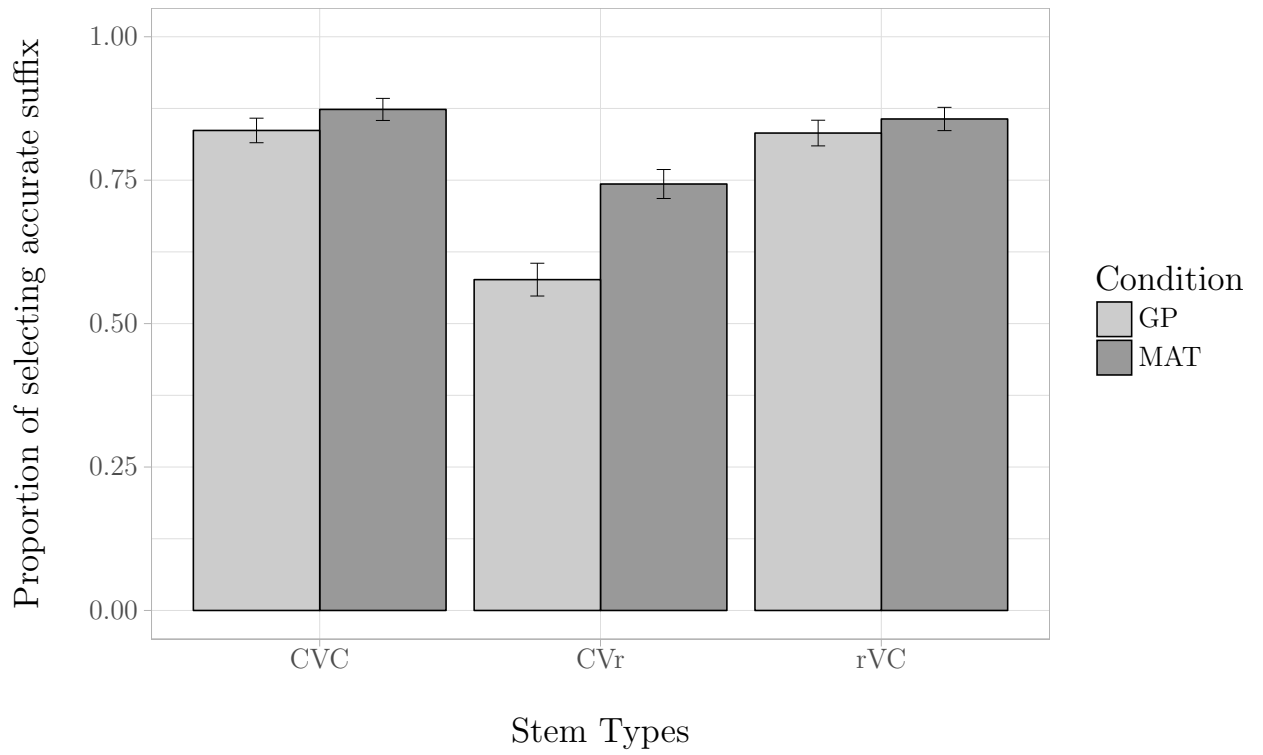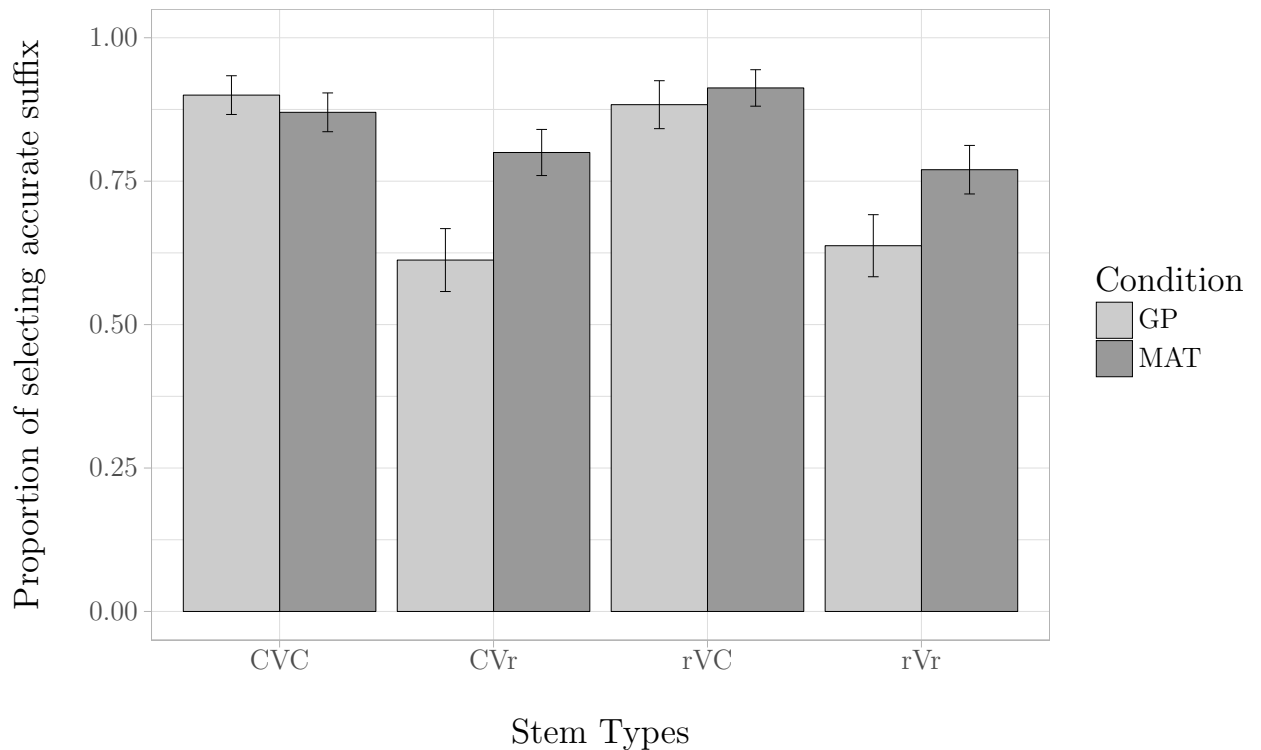

Figure 4.5: Proportion of selecting the conforming suffixes for the Old Items in Experiment 1A, with standard errors.
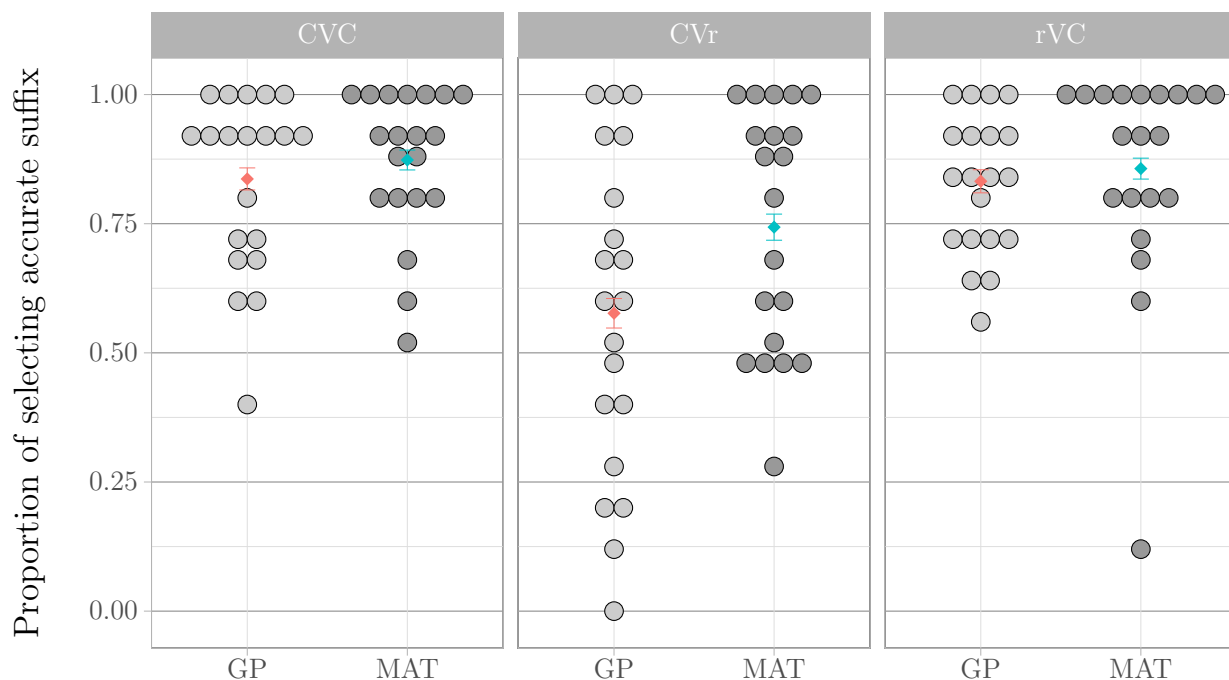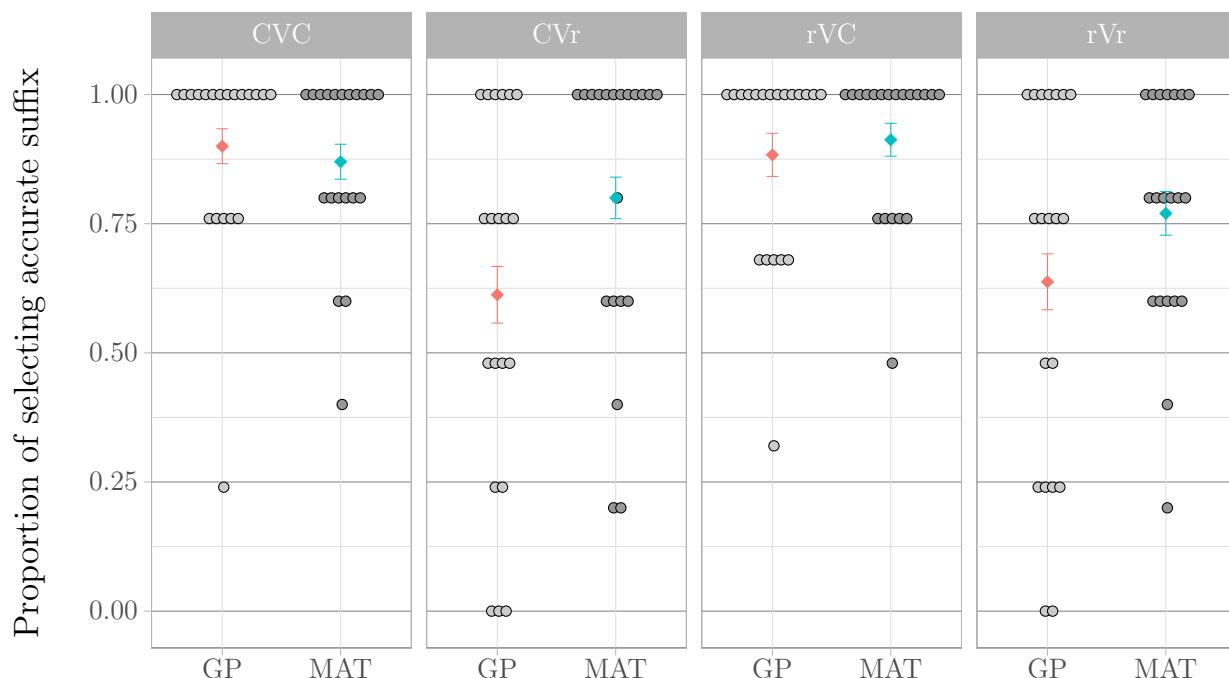
Figure 4.6: Proportion of selecting the grammatical suffix of each subject for Old Items in Experiment 1A. Each dot represents for a subject. Colored dots represent average with standard errors.

Table 4.5 shows the fixed effects for the Old Items. The baseline of the model shows the log-odds of choosing the suffix [-æt]. The negative intercept indicates that participants are less likely to choose the suffix. Among the stem types, CVr and rVr are significantly different from the baseline. The results show that there are no statistically significant differences in the interaction of the stem types and the condition ($p > 0.05$ in CVr×MAT, rVC×MAT, and rVr×MAT).

Table 4.5: Results of the fixed effects in the mixed-effects model for Old Items in Experiment 1A.

	Estimate	Std. Error	z value	Pr($>\mid z \mid$)	
(Intercept)	−2.482	0.437	−5.681	<0.001	***
CVr	3.016	0.478	6.313	<0.001	***
rVC	0.183	0.567	0.322	0.747	
rVr	3.140	0.481	6.526	<0.001	***
MAT	0.409	0.559	0.732	0.464	
CVr×MAT	0.582	0.621	0.938	0.348	
rVC×MAT	−0.643	0.760	−0.846	0.397	
rVr×MAT	0.266	0.618	0.431	0.666	

4.1.3 Discussion

The results of Experiment 1A show a difference in learning the CVr stem type. Specifically, participants were better at learning the CVr stem type in the MAT condition for the New Items. Other than the CVr stem type, the participants did not show any statistical differences in learning. CVr is the critical stem type because it shows whether participants actually learned the pattern. For the Old Items, there was no statistical difference between the two conditions.

MAT and the M-First model hypothesize that there will be no difference in learning the two grammars. Following the MAT-style analysis, both the GP grammar and the MAT grammar can be analyzed this way. To be more specific, one can categorize words into /r/-ending words which take /-æt/ in both conditions and other consonant ending words which take the /-et/ in the GP condition and /-ut/ in the MAT condition. There is no difference between the two conditions in terms of morpheme listing. The results show that participants do not just list or memorize the

71

morphemes in both conditions. Although both grammars could have been treated in the same way, the results show that participants do not treat the two grammars in the same way, and rejects the hypothesis. The M-FIRST model is also rejected because the model activates both morphological and phonological learning in both condition, which predicts no difference in learning the two grammars.

The P-FIRST model predicts that there will be a difference in learning in favor of phonologically motivated alternations. Following the model, the GP condition was predicted to be easier to learn since the suffix alternation has phonological motivation, i.e., vowel lowering. In other words, only the phonological learning is activated. This model predicts that the alternating pattern in the MAT condition is more difficult to learn because the alternation is phonologically unmotivated, and thus, has to activate both phonological learning and morphological learning. That is, the model treats the two grammars differently, and thus predicts a difference in learning results in favor of the GP condition. The results show that there is a difference in learning the two grammars, but in the opposite direction. This indicates that some factors other than morphological listing are reflected in the results.

4.1.3.1 Alternative Explanations

One of the possible factors for this outcome may be the phonetic saliency of the suffix. That is, the vowels in the alternant pair [-ut] and [-æt] are more phonetically distinct than the vowels in the [-et] and [-æt] pair. Although the [-ut] and [-æt] alternation is phonologically unmotivated, the phonetic salience between the vowels may have caused the alternation to be more discriminable than the [-et] and [-æt] pair, and thus it might have caught the attention of the participants in an experimental setting.

Discriminability (Blevins et al., 2016; Ackerman et al., 2016) may have facilitated the learning of the morpheme alternants. In a discriminative learning model, the grammar is learned more easily when the alternants are more discriminable (Ramscar et al., 2013). For example, as the irregular plural *mice/mouse* pair becomes discriminated from the regular plural *rats/rat*, the discrimination facilitates learning and thus

produces fewer overgeneralized forms (e.g., *mouses*) of the irregulars. If the more discriminable alternants accelerated learning of the grammar, it might be the case that participants in Experiment 1A learned the MAT condition better due to the discriminability.

By reducing discriminability between the vowels of the suffix alternants, we can conclude whether the learning differences are due to phonetic saliency. For this reason, Experiment 1B was set up the same way as Experiment 1A, except that the morpheme alternates between [-up] and [-un] in the MAT condition and [-em] and [-æm] in the GP condition. Again, the alternation between [-up] and [-un] is phonologically unmotivated whereas the same lowering process holds between [-em] and [-æm]. Therefore, it is more natural to analyze the pattern in the MAT condition as a phonologically-conditioned allomorphy with two underlying representations, whereas it is logical to analyze the pattern in the GP condition with phonology with a single underlying representation. Furthermore, the phonetic difference that existed between the vowels of the alternants in Experiment 1A is no longer present. Thus, the hypotheses still hold that people are predicted to learn the patterns in the GP condition better than the ones in the MAT condition.

4.2   Experiment 1B: [-em]∼[-æm] and [-up]∼[-un]

4.2.1   Methods

4.2.1.1   Subject

A total of 40 subjects (20 in each condition) participated in the experiment (mean age: 20.47; 11 were males). Subjects were recruited at the University of Delaware, were native speakers of English, and received either course credit or $5 as compensation.

4.2.1.2   Procedure

The procedure was the same as Experiment 1A.

4.2.1.3 Stimuli

The stimuli were the same as Experiment 1A, except for the corresponding suffixes. In the GP condition, participants were trained on the suffix [-em]∼[-æm] alternation, where the front vowels lower after a rhotic, with [-em] occurring elsewhere. In the MAT condition, participants were trained on the [-up]∼[-un] alternation, with [-un] occurring with stems ending with a rhotic, and [-up] occurring elsewhere. Table 4.6 shows the suffixes.

Table 4.6: Suffixes for the GP condition and the MAT condition in Experiment 1B.

Stem Types	Suffix	
	GP	MAT
CVC	-em	-up
CVr	-æm	-un
rVC	-em	-up
rVr	-æm	-un

All of the training and test items, and the number of items were the same as in Experiment 1A.

4.2.2 Results

The results show that participants were better at learning the CVr stem type in the New Items in the MAT condition. In the Old Items, participants were better at learning the CVr and the rVr stem type in the MAT condition.

Table 4.7 presents the proportion of the grammatical suffixes selected, with standard errors.

Table 4.7: Descriptive statistics of the rate of selecting the grammatical suffix in Experiment 1B with standard errors.

Stem Type	New		Old	
	GP	MAT	GP	MAT
CVC	0.780 (0.024)	0.850 (0.021)	0.825 (0.043)	0.910 (0.029)
CVr	0.583 (0.029)	0.737 (0.025)	0.625 (0.054)	0.820 (0.039)
rVC	0.803 (0.023)	0.897 (0.018)	0.713 (0.051)	0.880 (0.033)
rVr	–	–	0.563 (0.056)	0.790 (0.041)

The New Items and the Old Items were analyzed separately. Figure 4.7 shows the proportion of selecting the grammatical suffixes for the New Items and Figure 4.8 shows the proportion of selecting the grammatical suffixes for each subject. The colored dots represent the average of accuracy with standard error.

Figure 4.7: Proportion of selecting the grammatical suffix for the New Items in Experiment 1B, with standard errors.

Figure 4.8: Proportion of selecting the grammatical suffix of each subject for New Items in Experiment 1B. Each dot represents for a subject. Colored dots represent average with standard errors.

Results were analyzed with the same model that was used in Experiment 1A. Table 4.8 summarizes the fixed effects of the mixed-effects model for New Items in the experiment.

Table 4.8: Results of the fixed effects in the mixed-effects model for New Items in Experiment 1B.

	Estimate	Std. Error	z value	Pr $(>\mid z \mid)$
(Intercept)	-1.371	0.219	-6.265	<0.001 ***
CVr	1.731	0.191	9.082	<0.001 ***
rVC	-0.150	0.206	-0.726	0.468
MAT	-0.575	0.323	-1.780	0.075
CVr×MAT	1.377	0.298	34.615	<0.001 ***
rVC×MAT	-0.306	0.330	-0.928	0.353

The baseline of the model (intercept) is the log-odds of choosing the suffixes related to the rhotic-ending words (i.e., [-æm] in the GP condition and [-un] in the MAT condition). The negative intercept indicates that the subjects are less likely to choose the suffixes related to the rhotic-ending words for the CVC stem type in the GP condition. The log-odds of CVr indicate that the subjects are more likely to choose the suffixes in that stem type. The log-odds of rVC are not statistically different from the baseline. The log-odds of the condition (MAT) are not statistically significant. The interaction of the stem type and the condition indicates that subjects choose the conforming suffixes more often in the MAT condition with the CVr stem type ($p < 0.000$) but not with the rVC stem type ($p = 0.353$). This indicates that the difference between the GP condition and the MAT condition in the CVr stem type is statistically significant, and the participants were better at learning the pattern.

A one-sample t-test was conducted to determine whether the accuracy rate of the CVr stem type in the GP condition for the New Items ($M = 0.583$, $SD = 0.208$) was statistically different from a chance level, $t(19) = 12.515$, $p < 0.001$.

Figure 4.9 shows the proportion selecting the conforming suffixes for the Old

Items in the experiment and Figure 4.10 shows the proportion of selecting the grammatical suffixes for each subject. The colored dots represent the average of the accuracy rate with standard error.
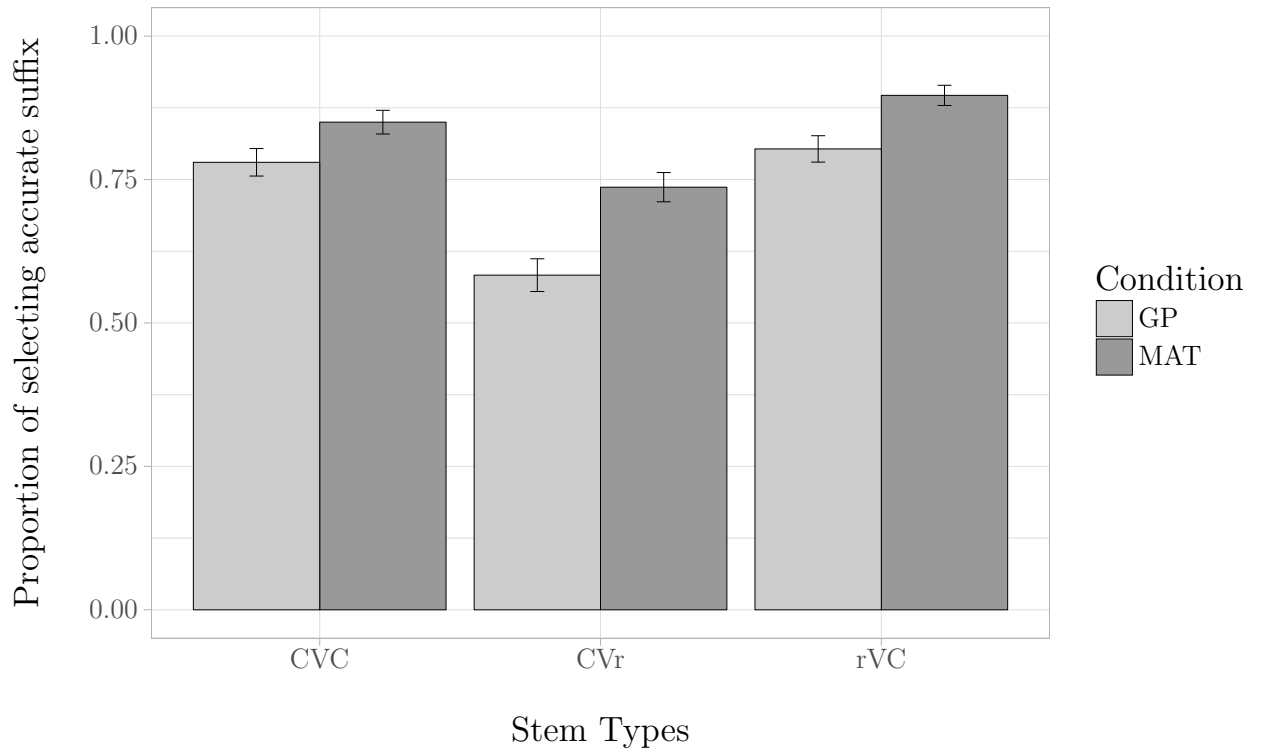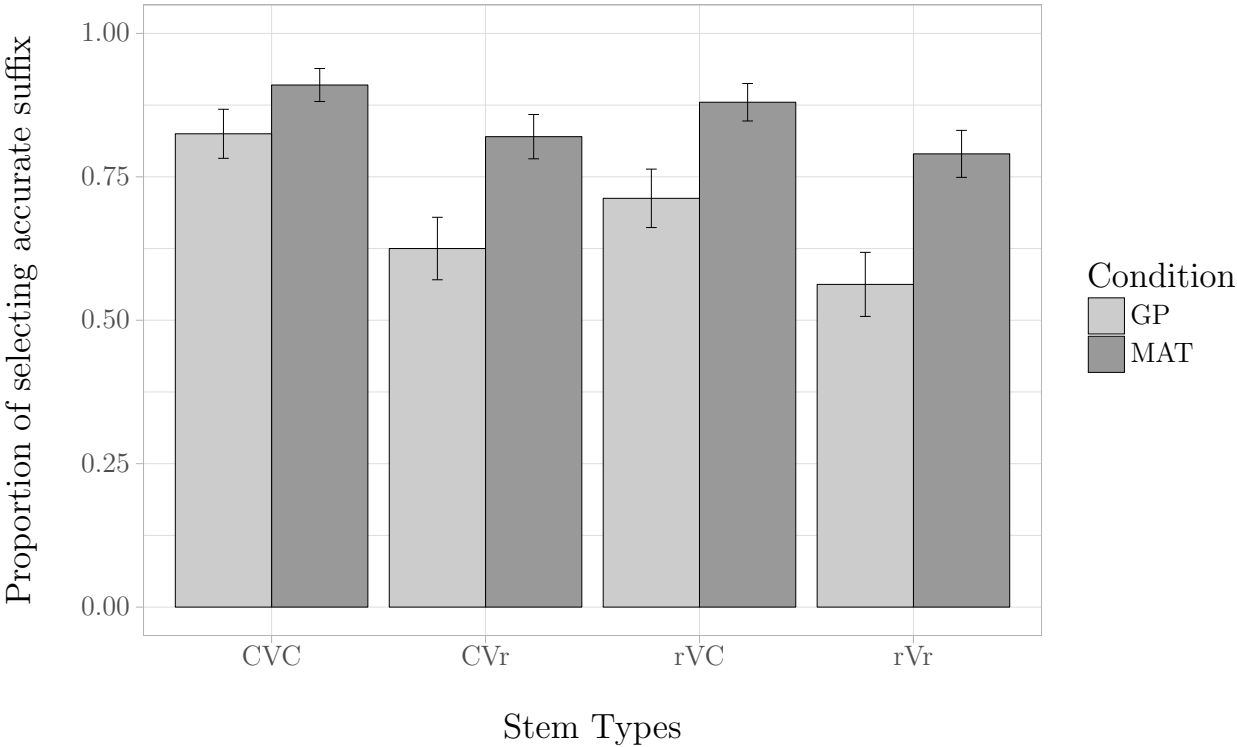

Figure 4.9: Proportion of selecting the grammatical suffixes for the Old Items in Experiment 1B, with standard errors.
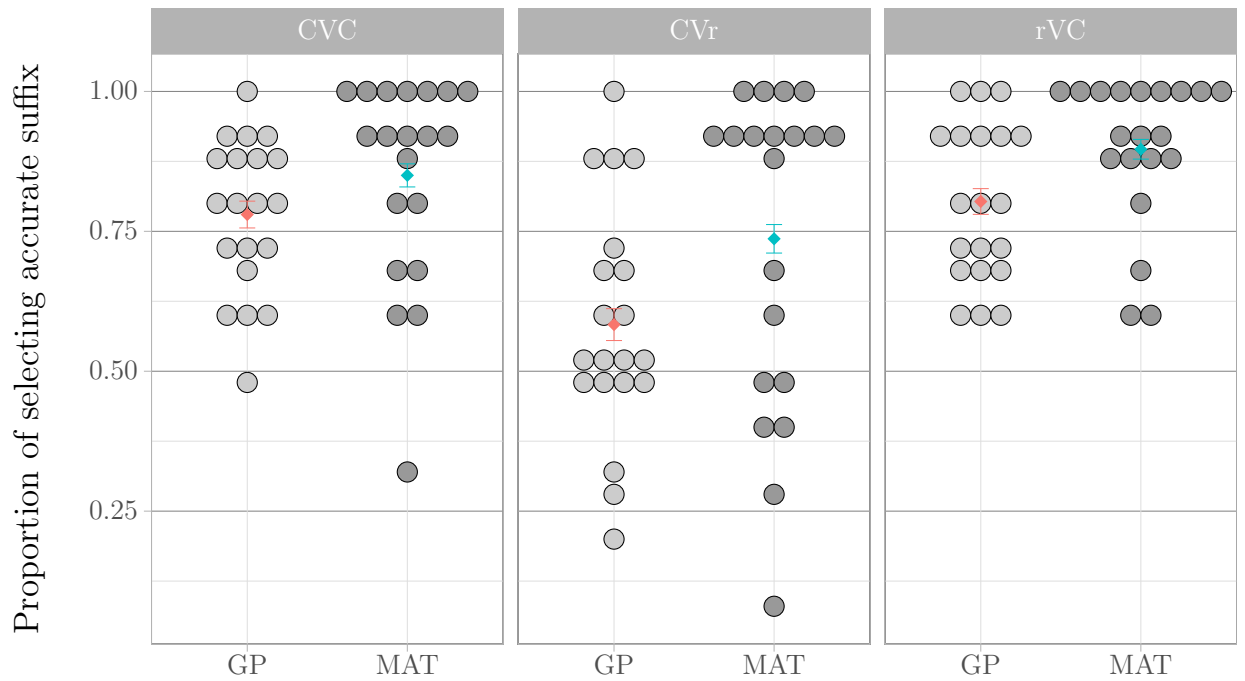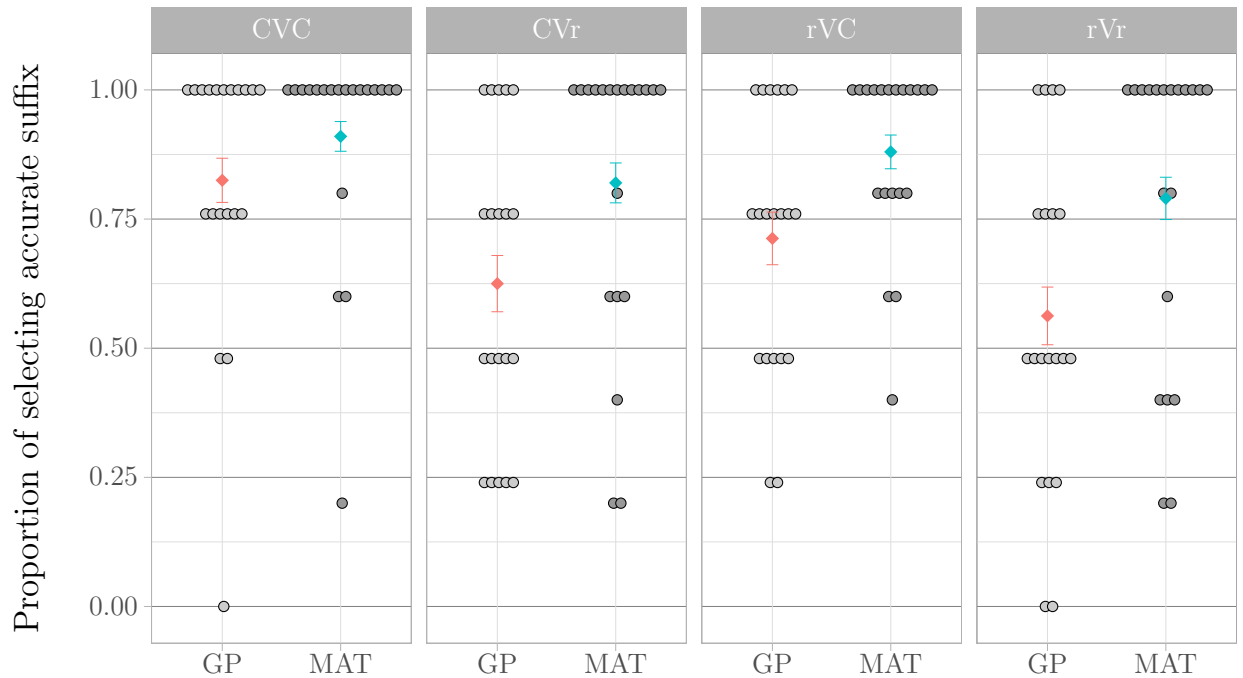
Figure 4.10: Proportion of selecting the grammatical suffix of each subject for Old Items in Experiment 1B. Each dot represents for a subject. Colored dots represent average with standard errors.

Table 4.9 shows the fixed effects for the Old Items. The baseline of the model shows the log-odds of choosing the suffixes related to a rhotic, and the negative intercept indicates that participants are less likely to choose those suffixes. Among the stem types, CVr and rVr are significantly different from the baseline. The interaction of the stem type and the condition indicates that the subjects choose the conforming suffixes more often in the MAT condition with the CVr ($p = 0.001$) and rVr ($p = 0.001$) stem types. This indicates that participants in the MAT condition were better at learning the CVr stem type and the rVr stem type than in the GP condition.

Table 4.9: Results of the fixed effects in the mixed-effects model for Old Items in Experiment 1B.

	Estimate	Std. Error	z value	Pr $(>|z|)$	
(Intercept)	-1.660	0.336	-4.946	<0.001	***
CVr	2.211	0.389	5.678	<0.001	***
rVC	0.680	0.394	1.723	0.085	
rVr	1.930	0.385	5.016	<0.001	***
MAT	-0.825	0.510	-1.617	0.106	
CVr×MAT	1.931	0.599	3.224	0.001	**
rVC×MAT	-0.348	0.616	-0.564	0.572	
rVr×MAT	2.004	0.590	3.400	0.001	***

4.2.3   Discussion

The results of Experiment 1B show a difference in learning the CVr stem type. Specifically, participants were better at learning the CVr stem type in the MAT condition for the New Items. Once again, CVr is the critical stem type because it shows whether participants learned the patterns. In the Old items, participants were better at learning the CVr and rVr stem type in the MAT condition.

For the New Items, the results were similar to Experiment 1A. MAT and the M-First model predicts that there should be no difference in learning the two grammars, but the results show that there was a difference in learning. The results of Experiment 1B also reject the MAT hypothesis and the M-First model. The P-First model predicts that the difference in learning will be in favor of the GP grammar, but the results show the opposite direction as Experiment 1A. Again, something other than morphological listing is going on.

4.2.3.1 Alternative Explanations

In the previous discussion, §4.1.3.1, one of the possible factors was the phonetic saliency between the vowels in the MAT condition. However, the results in Experiment 1B indicates that the better learning of the CVr stem type in the MAT condition in Experiment 1A was not due to the phonetic difference between [u] and [æ] compared to [e] and [æ] in the alternants.

A possible explanation for the learning results for this experiment is that the consonants [p] and [n] in the alternants [-up] and [-un] in the MAT condition are still discriminative. Consonants have been argued to play a role in identifying lexical items, whereas vowels play a role in abstracting structure (Hochmann et al., 2011; Toro et al., 2008). If this is the case, participants may discriminate the alternants in the MAT condition more easily, considering the alternants as a suppletive form, and thus participants may learn those alternants better than the ones in the GP condition. Therefore, participants may have learned the patterns in the MAT condition, including Old Items, better overall than those in the GP condition.

In addition to phonetic saliency, our memory system may have played a role when learning the patterns in each grammar. Declarative memory and procedural memory serve different roles when learning a language (the Declarative/Procedural Model, Ullman (2015)). Declarative memory learns the content of words, phonological forms of words, the meaning of words, and the irregular morphological forms (e.g., *fall - fell*). Procedural memory, in contrast, learns implicit rules, which are very important in "learning to predict" (Ullman, 2015, p.960) a grammar that results from linguistic rules. The two memory systems interact when learning a language; declarative memory begins to learning from the earlier stage, while the procedural memory gradually learns the complex rules.

If we apply the declarative/procedural model to learning the GP grammar and the MAT grammar, we can speculate why people were better at learning the allomorph patterns in the MAT condition in Experiments 1A and 1B. If the participants listed or mapped each alternant in the MAT condition to its own underlying representation, it

is possible that the alternants were treated as suppletive irregular forms that depend on declarative memory. In contrast to learning in the MAT condition, if the participants in the GP condition mapped alternants to a single underlying representation as well as abstracting rules for phonological transformation, they would depend more on procedural memory. Since declarative memory begins learning at an earlier stage than procedural memory, we can speculate this as the reason why the participants in the MAT condition learned the allomorphy pattern better than the participants in the GP condition. In other words, the two memory systems may explain why people learn one pattern faster than the other pattern.

The memory systems lead to a question regarding how quickly people know that the patterns in each grammar are different in terms of lexical representation and types of rules. In other words, within a short amount of time or exposure to the patterns, how do people know that the alternants in the MAT condition are suppletive but the alternants in the GP condition form phonological transformations? Why do people treat the patterns in the GP condition differently from the patterns in the MAT condition?

Lastly, the participants could set up hypotheses about the patterns when learning. Gerken and Knight (2015) studied infants who were able to make phonological generalizations in an artificial grammar learning experiment, when the patterns can be generalized into a non-conflicting rule and a partially-conflicting rule. That is, in a non-conflicting rule condition, patterns were only generalized with a feature-based rule, whereas in a partially-conflicting rule condition, all patterns were generalized with a feature-based rule and some patterns were also generalized with consonant-based rules. However, when there were two competing rules for generalizing the pattern, infants failed to make the generalization. In other words, it is more difficult to learn a pattern when the pattern could be analyzed in more than one way. For example, in the MAT condition, a learner may make a generalization such as "When a word ends with [r], select the [-un] suffix, and [-up] otherwise." In the GP condition, however, there could be two possible analyses that a learner could make. First, one could generalize in a

similar way as the learner in the MAT condition: "When a word ends with [r], select the [-æm] suffix, and [-em] otherwise." Second, one could generalize the implicit phonological rules, i.e., the vowel lowering process after the rhotic. In other words, in the GP condition, there are two competing generalizations that the participants could make, which might have made the learning more difficult. As is pointed out in the article, people may need more information to make a broader or "winning" generalization.

Therefore, in Experiment 2, I include both of the suffixes used in Experiments 1A and 1B so that the participants are exposed to a stronger generalization for the phonological transformation. Participants in the GP condition learn the morpheme alternant pairs [-et]∼[-æt] and [-em]∼[-æm]. The two suffixes show the same phonological transformation of vowel lowering, and thus, it is predicted that it will be easier to for the participants to generalize the phonological transformation. In the MAT condition, participants learn [-ut]∼[-æt] and [-up]∼[-un]. The suffixes do not share any phonological transformation, and thus, all four forms should be listed in the lexicon. Therefore, it is predicted that it will be more difficult for the participants to generalize the pattern. Experiment 2 is presented in next chapter.

Chapter 5

EXPERIMENT 2

This chapter presents Experiment 2, in which the participants learned two pairs of suffixes in each condition. The two pairs of suffixes are the same morphemes that were used in Experiments 1A and 1B.

5.1 Methods

5.1.1 Subject

A total of 80 subjects (40 in each condition) participated in the experiment (mean age: 21.0; and 9 were males). Within each condition, there were two versions; 20 subjects participated in each version. The participants were recruited at the University of Delaware, were native speakers of English, and they received either course credits or $10 as compensation.

5.1.2 Procedure

The overall procedure was the same as in the previous experiments, except that the training session and the test session were longer. The sessions were longer because participants were trained and tested on two suffixes. The training session had two subsessions. Participants learned one pair of suffixes during the first subsession of the training, and then learned the other pair during the second subsession. Before the actual experiment, participants were told that there will be two training sessions one after another, and one test session. In the experiment, participants read the instruction on the screen that they will learn the plural suffix in the first training session and then they were exposed to the training. After the first training session, they moved on to the second training session and they read the instructions on the screen that they will

learn the possessive suffix. The total duration of the experiment was approximately 45 to 50 minutes (training approximately 25 minutes and test approximately 20 to 25 minutes) without any break.

5.1.3 Stimuli

Experiment 2 included all of the allomorphs from Experiments 1A and 1B in both conditions. Thus, there were two versions according to the order in which the suffixes were presented. In Version 1 of the GP condition, participants were first trained on the [-em]∼[-æm] pair and then on the [-et]∼[-æt] pair. In Version 2, participants were first trained on the [-et]∼[-æt] pair and then on the [-em]∼[-æm] pair. In Version 1 of the MAT condition, participants were first trained on the [-up]∼[-un] pair, and in Version 2 they were trained first on the [-ut]∼[-æt] pair.

Table 5.1: Suffixes for the GP condition and the MAT condition in Experiment 2. In each version, the first form was presented before the second one.

Stem Types	Version 1		Version 2	
	GP	MAT	GP	MAT
CVC	-em/-et	-up/-ut	-et/-em	-ut/-up
CVr	-æm/-æt	-un/-æt	-æt/-æm	-æt/-un
rVC	-em/-et	-up/-ut	-et/-em	-ut/-up
rVr	-æm/æt	-un/æt	-æt/-æm	-æt/-un

The number of training items and test items were in Experiment 2 doubled the number used in Experiments 1A and 1B. All of the training and the test stimuli were the same as in Experiments 1A and 1B.

5.2 Results

The results show that participants learned the CVr stem type in New Items in the MAT condition better, but no difference in the Old Items.

Table 5.2 shows the proportion of times participants selected the grammatical suffix. The grammatical suffixes for CVC and rVC stem types are [-et] and [-em] in the GP condition and [-ut] and [-up] in the MAT condition. The grammatical suffixes for CVr and rVr stem types are [-æt] and [-æm] in the GP condition and [-æt] and [-un] in the MAT condition.

Table 5.2: Descriptive statistics of rate of selecting the grammatical suffix in Experiment 2, with standard errors in parentheses.

Stem Type	New		Old	
	GP	MAT	GP	MAT
CVC	0.791 (0.012)	0.798 (0.012)	0.825 (0.021)	0.823 (0.019)
CVr	0.537 (0.014)	0.589 (0.014)	0.634 (0.027)	0.648 (0.024)
rVC	0.794 (0.012)	0.788 (0.012)	0.797 (0.023)	0.820 (0.019)
rVr	–	–	0.597 (0.027)	0.620 (0.024)

Figure 5.2 shows the proportion of selecting the grammatical suffixes in the New Items and Figure 5.2 shows the proportion of selecting the grammatical suffixes for each subject. The colored dots represent the average with standard error.

Figure 5.1: Proportion of selecting the grammatical suffix for the New Items in the Experiment 2, with standard errors.

Figure 5.2: Proportion of selecting the grammatical suffix of each subject for New Items in Experiment 2. Each dot represents for a subject. Colored dots represent the average with standard errors.

Results were analyzed with a mixed-effect model with a binomial function using the *glmer function included in the* lme4 package (Bates et al., 2015) in *R* (R Core Team, 2017).

The subjects' responses were coded as '1' if they chose the grammatical suffix of the rhotic-ending words (i.e., the suffixes [-æt] and [-æm] in the GP condition and [-æt] and [-un] in the MAT condition) and '0' otherwise. The CVC stem type in the GP condition was the reference level. Table 5.3 summarizes the fixed effects of the mixed-effects model for the New Items in the experiment.

Table 5.3: Results of the fixed effects in the mixed effects model for New Items in Experiment 2.

	Estimate	Std. Error	z value	Pr $(>\lvert z \rvert)$	
(Intercept)	-1.402	0.128	-10.944	<0.001	***
CVr	1.567	0.102	15.363	<0.001	***
rVC	-0.026	0.110	-0.237	0.812	
MAT	-0.117	0.178	-0.656	0.512	
CVr×MAT	0.338	0.136	2.480	0.013	*
rVC×MAT	0.092	0.148	0.622	0.534	

The negative intercept of the log-odds indicates that the subjects are less likely to choose the suffixes [-æt] and [-æm] in the GP condition and [-æt] and [-un] in the MAT condition. The log-odds of CVr indicate that the subjects are more likely to choose the suffixes related to the rhotic in the stem type. The log-odds of rVC indicate that subjects are less likely to choose the suffix not related to rhotics. The log-odds of condition are not statistically significant ($p = 0.512$). The interaction of the stem type and the condition indicates that subjects choose the conforming suffixes more in the MAT condition with the CVr stem type ($p = 0.013$) but not with the rVC stem type ($p = 0.534$). Again, this indicates that participants were better at learning the CVr stem type in the MAT condition than in the GP condition in the New Items.

A one-sample t-test was conducted to determine whether the accuracy rate of the CVr stem type for the New items in the GP condition ($M = 0.537$, $SD = 0.253$) was statistically different from a chance level, $t(39) = -1235.2$, $p < 0.001$. The same test was conducted to determine whether the accuracy rate of the CVr stem type for the New items in the MAT condition ($M = 0.589$, $SD = 0.261$) was statistically different from a chance level, $t(39) = -1197.4$, $p < 0.001$.

Figure 5.3 shows the proportion selecting the conforming suffixes for Old Items in the Experiment 2 and Figure 5.4 shows the proportion selecting the grammatical suffixes for each subject. The colored dots represent the average of the accuracy rate with standard error.
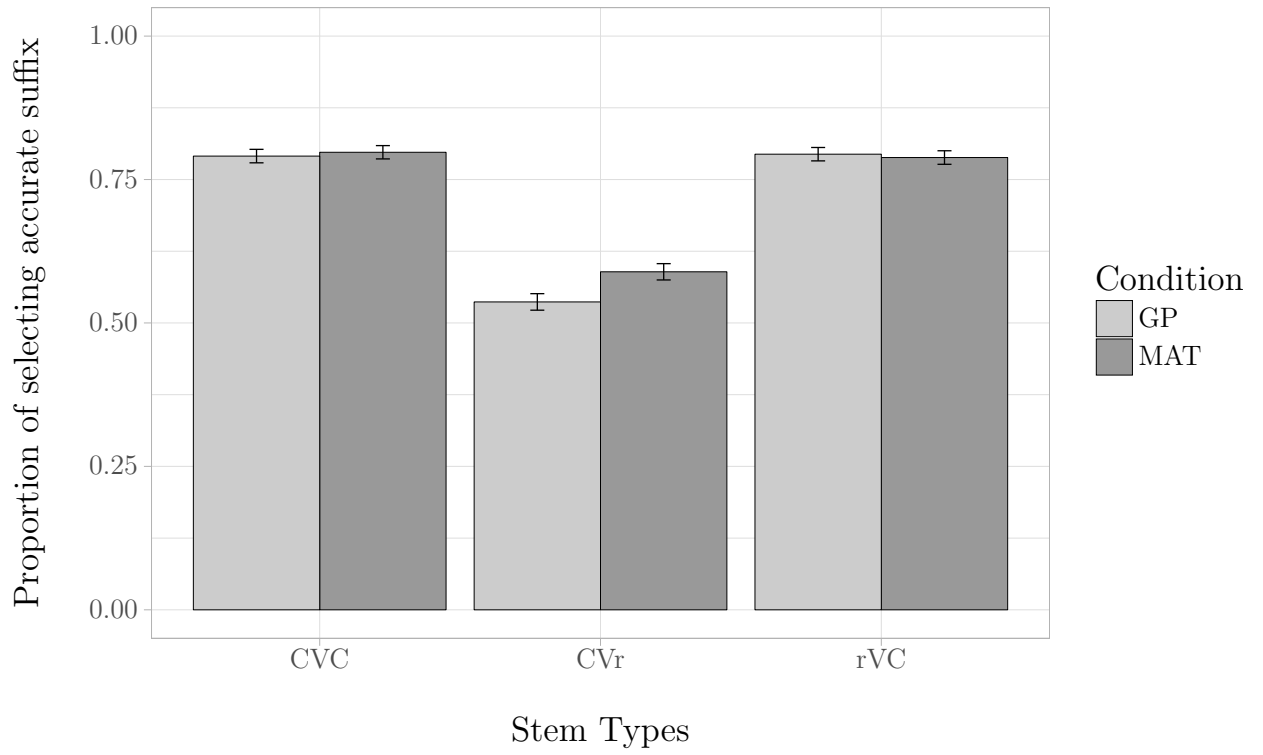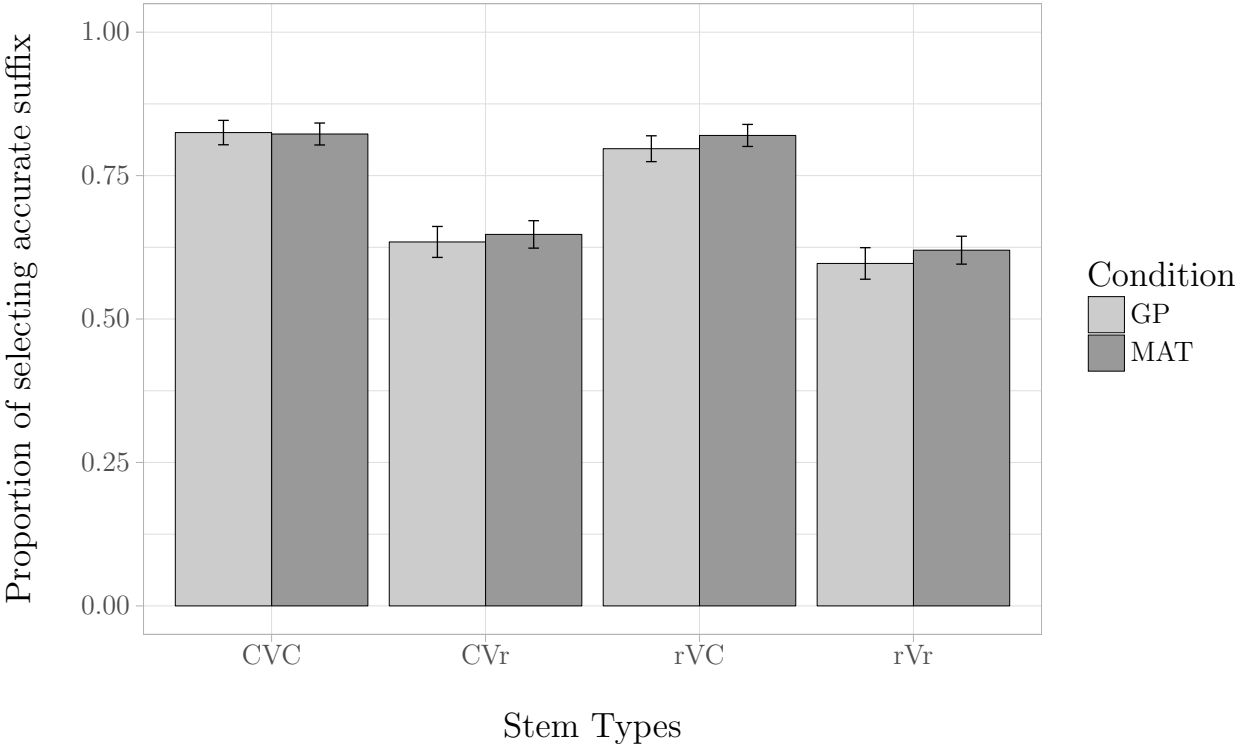

Figure 5.3: Proportion of selecting the grammatical suffixes for the Old Items in Experiment 2, with standard errors.
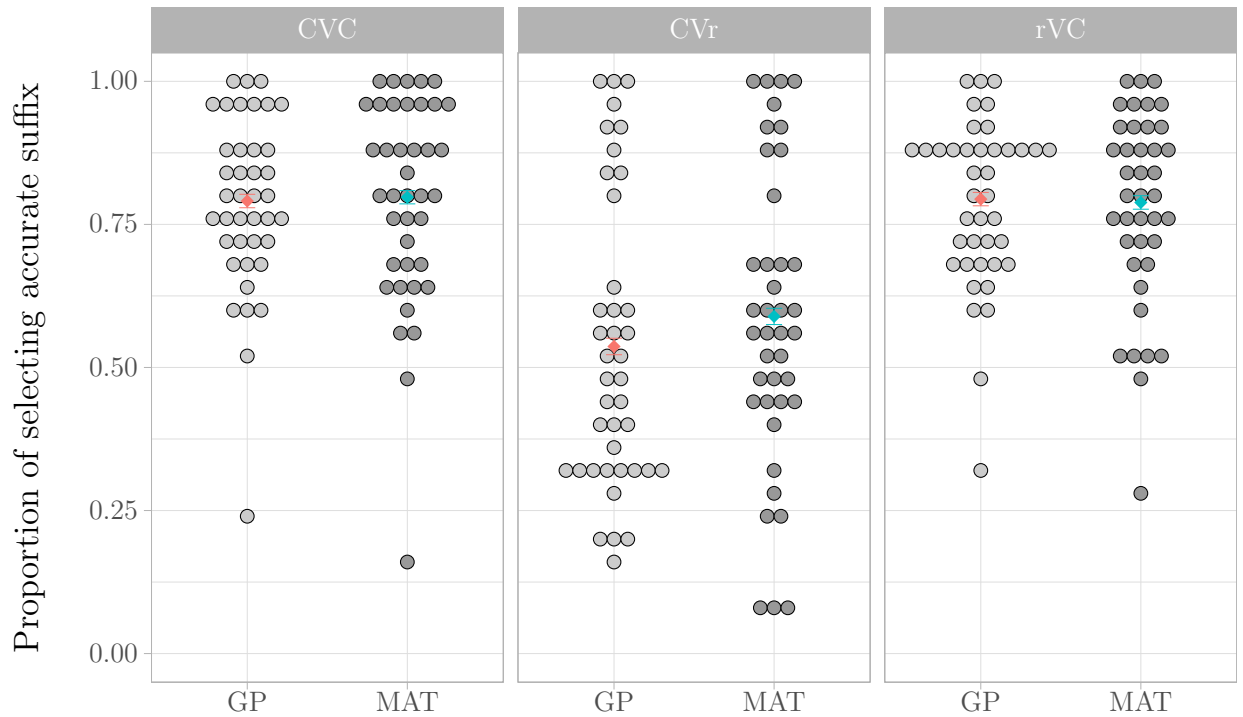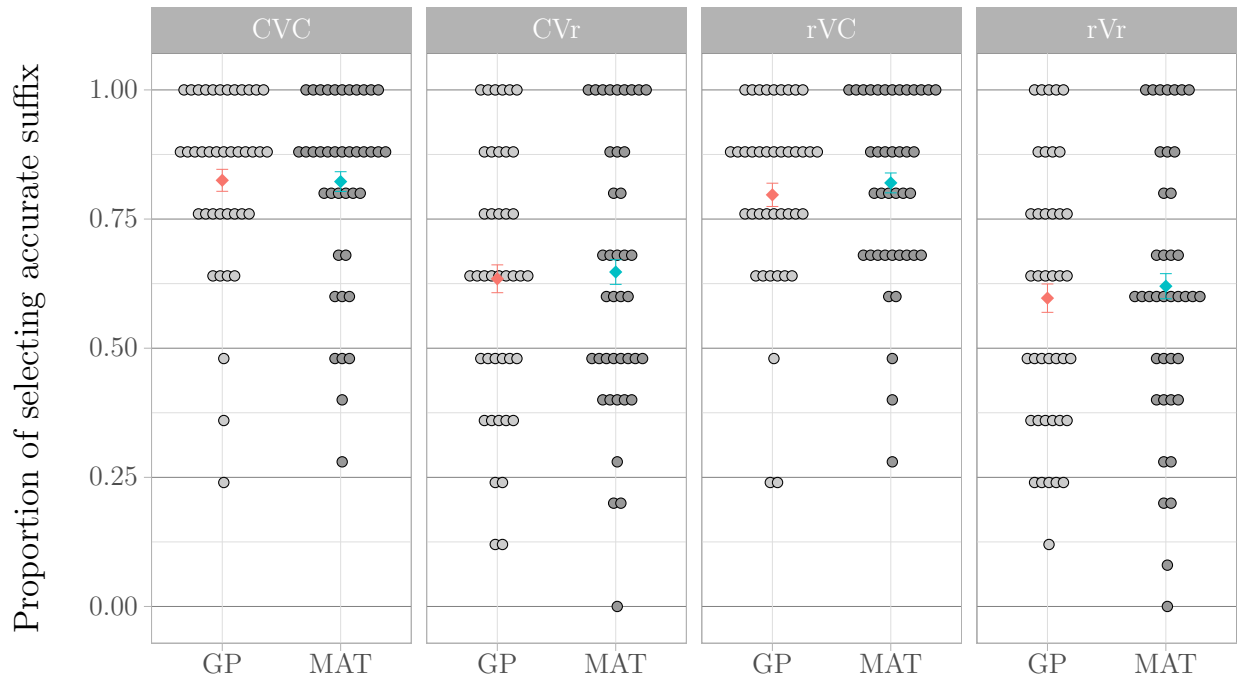
Figure 5.4: Proportion of selecting the grammatical suffix of each subject for Old Items in Experiment 2. Each dot represents for a subject. Colored dots represent the average with standard errors.

Table 5.4 shows the fixed effects for the old items. The negative intercept of the baseline indicates that participants were less likely to choose the suffix. Among the stem types, CVr and rVr are significantly different from the baseline. The results show there is no statistically significant difference in the interaction of the stem types and the condition ($p > 0.05$ in CVr×MAT, rVC×MAT, and rVr×MAT). This indicates that there was no difference in learning the GP condition and the MAT condition.

Table 5.4: Results of the fixed effects in the mixed-effects model for Old Items in Experiment 2.

	Estimate	Std. Error	z value	Pr $(>|z|)$	
(Intercept)	-1.666	0.200	-8.312	<0.001	***
CVr	2.262	0.227	9.945	<0.001	***
rVC	0.188	0.238	0.788	0.430	
rVr	2.062	0.226	9.110	<0.001	***
MAT	-0.044	0.256	-0.173	0.863	
CVr×MAT4	0.107	0.271	0.394	0.694	
rVC×MAT4	-0.138	0.289	-0.477	0.633	
rVr×MAT4	0.175	0.272	0.645	0.519	

5.2.1 Additional Analysis

Additional analyses were conducted to see the differences across the experiments. Figure 5.5 shows the proportion selecting the grammatical suffix for the New Items in each experiment.
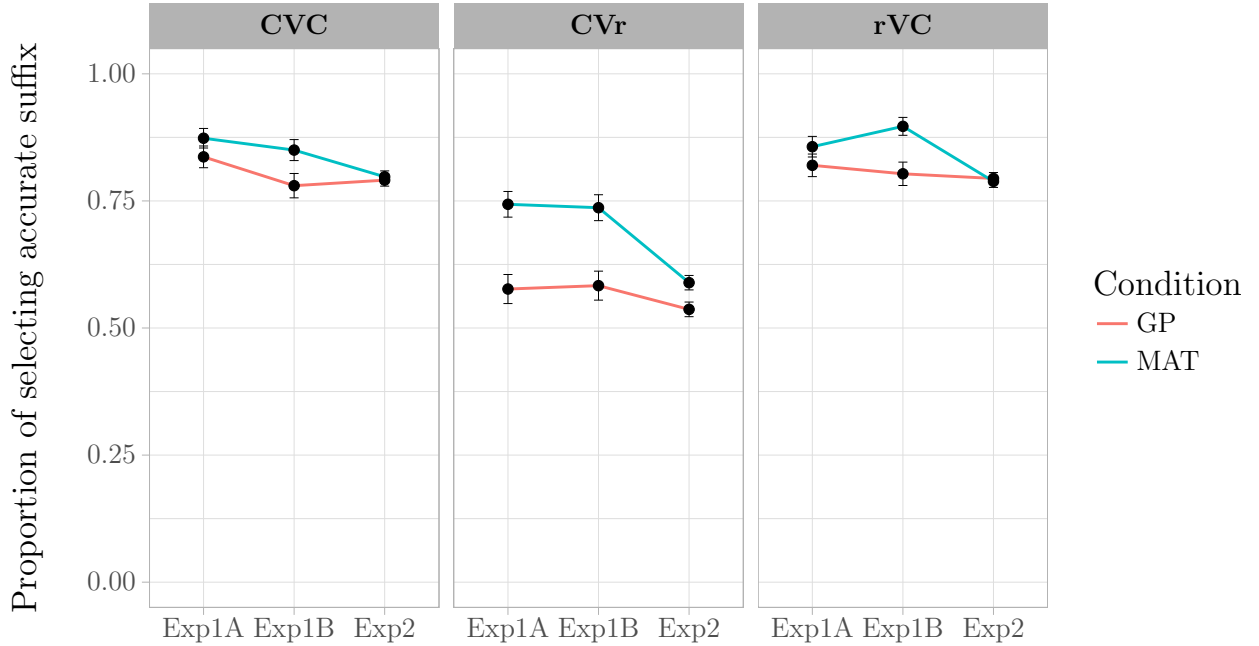
Figure 5.5: Proportion of selecting the grammatical suffix for the New Items in each experiment, with standard errors.

One-way ANOVA was conducted to compare the mean differences between experiments within the condition and the stem type. In the CVC stem type, there was no statistical significance between experiments in the GP condition ($F(2, 77) = 0.786$, $p = 0.459$) and in the MAT condition ($F(2, 77) = 1.482, p = 0.234$). In the CVr stem type, there was no statistical significance between experiments in the GP condition ($F(2, 77) = 0.282, p = 0.755$), whereas there was statistical significance between the experiments in the MAT condition ($F(2, 77) = 3.257, p = 0.044$). Post-hoc pairwise t-test showed that accuracy rate of Experiment 2 is significantly lower than Experiment 1A ($p = 0.036$) and Experiment 1B ($p = 0.045$) in the MAT condition. This indicates that participants in Experiment 2 in the MAT condition were worse at

learning the patterns than the participants in Experiments 1A and 1B in the MAT condition. In the rVC stem type, there was no statistical significance between experiments in the GP condition ($F(2, 77) = 0.476, p = 0.623$) and in the MAT condition ($F(2, 77) = 2.887, p = 0.0618$).

Figure 5.6 show the proportion of selecting the grammatical suffix for the Old Items in each experiment.
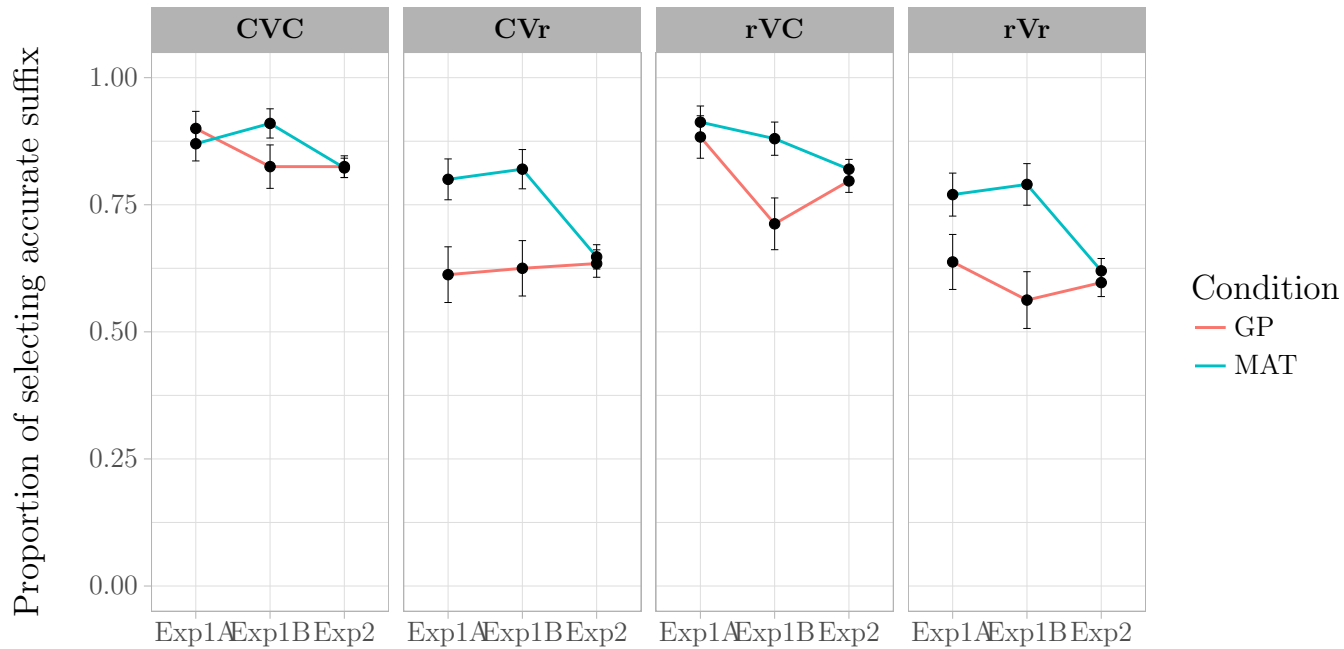
Figure 5.6: Proportion of selecting the grammatical suffix for the old Items in each experiment, with standard errors.

One-way ANOVA was conducted to compare the mean differences between experiments within the condition and the stem type. In the CVC stem type, there was no statistical significance in the GP condition ($F(2, 77) = 1.021, p = 0.365$) and in the MAT condition ($F(2, 77) = 1.478, p = 0.234$). In the CVr stem type, there was

no statistical significance in the GP condition ($F(2, 77) = 0.039, p = 0.962$), whereas there was statistical significance in the MAT condition ($F(2, 77) = 3.513, p = 0.035$). Post-hoc pairwise t-test showed that accuracy rate of Experiment 2 was significantly lower than Experiment 1A ($p = 0.047$) and Experiment 1B ($p = 0.025$) in the MAT condition. This indicates that the participants in Experiment 2 were worse at learning the patterns than the participants in Experiments 1A and 1B in the MAT condition. In the rVC stem type, there was a statistical significance in the GP condition ($F(2, 77) = 3.488, p = 0.036$), whereas there was no statistical significance in the MAT condition ($F(2, 77) = 2.175, p = 0.121$). Post-hoc pairwise t-test showed that the accuracy rate of Experiment 1B was significantly lower than Experiment 1A ($p = 0.01$) but not significantly lower than Experiment 2 ($p = 0.14$). This means that participants in Experiment 1B in the GP condition were worse at learning the rVC stem type than in Experiments 1A and 2. In the rVr stem type, there was no statistical significance in the GP condition ($F(2, 77) = 0.322, p = 0.726$), whereas there was statistical significance in the MAT condition ($F(2, 77) = 3.609, p = 0.032$). Post-hoc pairwise t-test showed that the accuracy rate of Experiment 2 was significantly lower than Experiment 1A ($p = 0.44$) and Experiment 1B ($p = 0.023$). This indicates that participants in Experiment 2 were worse than the participants in Experiments 1A and 1B at learning the rVr stem type in the MAT condition.

5.3 Discussion

The results of Experiment 2 show a difference in learning: the participants were better at learning the CVr stem type in the MAT condition for the New Items. There was no statistical difference in learning within the Old Items. The learning results show the same tendency as Experiments 1A and 1B.

The results were similar to Experiments 1A and 1B in that the results reject the MAT hypothesis and the M-FIRST model. Across the experiments, participants did not treat the two grammars in the same way. However, the results do not fully

support the P-First model, either. The P-First model predicts that there will be a difference in learning, especially in favor of the GP grammar.

Although the question why the CVr stem type was learned better in the MAT condition remains, the results of the experiments show an interesting tendency. Despite the fact that participants in the MAT condition were better at learning the CVr stem type across all experiments, the rate of selecting the grammatical suffix decreased significantly, from 74.3% and 73.7% in Experiments 1A and 1B respectively to 59% in Experiment 2. It seems that when there are more morphemes to memorize or list in the lexicon, participants have more difficulty learning or memorizing them. Thus, one must question whether the tendency would still hold when there are more morphemes to learn. It is expected that the tendency will hold, in other words, participants are expected to have more difficulty learning the MAT grammar than the GP grammar with greater number of morphemes. If the tendency still holds, then, it fully supports the P-First model.

It is important to study this topic because it tells us the relationship between the cognitive load of learning or memorizing morphemes and the cognitive load of learning phonological transformational rules. In addition, it is important to ask how much information one needs as a learner to make broader or "winning" generalizations (Gerken and Knight, 2015), and finally learn the phonology of the language. Importantly, the results reject MAT and the M-First model even with one pair of morpheme alternants, indicating something more than morpheme listing is going on when learning the alternants. The results in this dissertation implicate that two pairs of morpheme alternants were not enough to fully support the P-First model. Then, the ensuing question is how much information one needs to make a phonological generalization?

5.3.1 Alternative Explanations

Related to the number of the morphemes to learn, it seems that the role of discriminability or phonetic saliency seems to weaken when the number of morphemes

to learn increases. In Experiment 2, the phonetic saliency between vowels and consonants in the MAT condition (i.e., [u]∼[æ] and [p]∼[n]) was not as crucial as it was in Experiments 1A and 1B. In other words, in the GP condition, the differences were only on the vowels [e]∼[æ], but in the MAT condition, the differences were in both vowels and consonants. Therefore, if phonetic saliency plays a critical role in learning the morphemes, participants are expected to learn the morphemes in the MAT condition better, and the participants in the MAT condition were better at learning the CVr stem type. Yet, the difference between the MAT condition and the GP condition is smaller in Experiment 2 than in Experiments 1A and 1B. This is interesting because if phonetic saliency plays an important role in learning the allomorphs, participants in Experiment 2 should perform better than the participants in Experiments 1A and 1B in the MAT condition. In the MAT condition, more phonemes are distinctive compared to the allomorphs in the GP condition (i.e., both [u]∼[æ] and [p]∼[n] in the MAT condition and only [e]∼[æ] in the GP condition). Therefore, if phonetic saliency plays a crucial role, participants are expected to learn the allomorphs better in Experiment 2 in the MAT condition; but this is not the case.

To confirm the role of phonetic saliency in learning the morphemes, one should compare them to the allomorphs with a minimal consonantal difference such as [-up]∼[-ut] for the MAT condition. The alternation is less salient than [p]∼[n] because the alternation [p]∼[t] are in the same natural class which only differ in place of articulation. Since the alternation still lacks phonological motivations, therefore, the phonologically-conditioned selection rules could be, for example, 'select [-up] when the word ends with /r/, and [-ut] elsewhere'. If everything including the phonetic saliency is equal, and if people successfully learn the allomorphy pattern with the minimal phonetic difference as in Experiments 1A and 1B, then this suggests that learning involves other cognitive processes. If all things being equal, and if people have more difficulty to learn the allomorphy pattern than Experiments 1A and 1B, then this suggests that the learning result in the MAT condition was due to phonetic saliency.

To summarize Experiments 1A, 1B, and 2, the results consistently show that

participants were better at learning the CVr stem type in the MAT condition than in the GP condition. The results reject MAT and the M-FIRST model consistently in that participants treat the two grammars differently. The results partially support the P-FIRST model in the opposite direction. It is expected that with a greater number of morphemes, the the P-FIRST model would be fully supported. These results raise many interesting questions, and thus can prompt many follow-up studies. Possible follow-up studies, including their research questions, are discussed in the following chapter.

Chapter 6

GENERAL DISCUSSION

This dissertation investigated the esearch question whether humans prefer morphological listing or phonological analyses when learning morpho-phonological patterns that vary in their pronunciations. A series of AGL experiments were conducted to examine whether there is any difference in learning that would favor either theory.

The morpho-phonological patterns in the artificial grammars support each theoretical view. The alternant pattern in the GP grammar was based on phonological transformation and thus the allomorphs are more likely to be analyzed as a single underlying representation. However, the GP grammar can be also analyzed with morpheme listing as in MAT by categorizing the words into groups that share the same phonological ending. The alternant pattern in the MAT grammar was based on phonologically-conditioned allomorphy, and all of the alternants are more likely to be listed on their own. Unlike the GP grammar, the MAT grammar is less likely to be analyzed with phonological transformation as in GP.

The null hypothesis was that there will be no difference in learning and the experimental hypothesis was that there will be a difference in learning the two grammars. The null hypothesis was based on MAT and the M-First model. The experimental hypothesis was based on the P-First model. The model predicts that it will be easier to learn the GP grammar. In this model, phonological learning is activated before morphological learning, and therefore, one would learn the phonologically analyzable patterns better.

The consistent finding that participants in the MAT condition learned the CVr stem type better across all experiments rejects MAT and the M-First model, and

100

weakly supports the P-First model because the difference was in the opposite direction. The unexpected results prompt the discussion issues of what accounts for the observed learning differences.

The first issue focuses on the learning strategy or the cognitive process of learning the allomorphs. The Simultaneous model was not discussed in this dissertation because it is difficult to know which modules play a role and compete when learning a language. If we consider the cognitive process or cognitive load as one of the modules that is activated, the results raise an interesting question. In the MAT condition, people had to learn the four underlying representations, /-ut/, /-æt/, /-up/, and /-un/, as the lexical representations, in addition to the phonologically-conditioned selection rules, categorizing the stem types into those that end with /r/ and others. In the GP-condition, people had to learn two underlying representations, /-et/ and /-em/, and a phonological transformation, vowel lowering. If we only consider the number of representations and the number of rules to learn, MAT should be more cognitively demanding. However, the results show that people were better at learning the MAT-grammar. If we assume that memorizing or listing all of the alternants along with the phonologically-conditioned selection rules is easier or requires less cognitive load than abstracting the underlying representation for the alternants and phonological transformations, then the question is why did the participants in the GP-condition not memorize the pronunciations? In other words, if people chose to memorize the alternants in the MAT condition, they could have memorized them in the GP-condition as well.

This question speaks to anecdotal observations that I came across during the study. Participants in the MAT condition were more likely to identify the suffix alternants as "selecting a different suffix when a word ends with [r]" after the experiment. In the GP condition, however, the responses were mixed. The Participants who reported the morpheme alternants as "select [-æ] suffix when a word ends with [r]" showed a high accuracy rate most of the time. Other participants, who reported that they knew there were two pronunciations of each suffix but were not sure which alternant goes

with which stem type, showed mixed results for accuracy. That is, some of them showed a high accuracy rate even though they were not able to describe or explain what the pattern was, while some of them consistently chose the [-et] and [-em] alternants most of the time. It is interesting that many participants could easily identify the allomorphs with the selection rule in the MAT condition, while fewer people analyzed the suffixes that way in the GP condition.

A second issue for the results may be a task effect. The two-alternative forced choice (2FAC) task is a widely used testing task in the AGL paradigm (Finley, 2015; Lai, 2015, and many others), and show the learning results. However, the two-alternative forced choice (2FAC) task during the testing phase of the experiments may not have been the best method for testing the learnability of the allomorphs. Peperkamp and Dupoux (2007) studied whether natural classes play a role when learning allophones by using a forced-choice phrase-picture-matching task in the test phase of an artificial grammar learning paradigm. Their results showed that participants learned the allophones in the phonologically natural class just as well as the ones in the phonologically unnatural class. In Peperkamp et al. (2006), they used the same training and test stimuli as Peperkamp and Dupoux (2007) but changed the testing method to a picture-naming task. They thought the forced-choice phrase-picture-matching task might have been too easy of a task or too explicit, and thus the participants may have used other, non-linguistic, learning methods to learn the unnatural class. With the picture-naming task, participants had more difficulty producing the allophones in the unnatural class. This raises an interesting question about the testing method in the experiments.

In this dissertation, participants were given a word-suffix pair in which one of the suffixes was the correct form and the other was not. That is, participants heard both correct and incorrect forms back to back and were asked to choose one of the options that they thought was the right form. In the two-alternative forced choice task, it is possible that the task might have contrasted the correct and incorrect form, and thus, facilitating selection of the correct answer by comparison with the correct and

incorrect form from the test phase. However, this does not explain why this contrast facilitated the comparison more in the MAT condition than in the GP condition. The easiest way to overcome the possibility of contrast in a two-alternative forced choice task is to change the test method to a 'yes-no' task. The 'yes-no' task is still forced-choice, but only one of the correct or incorrect form is presented to the participants. If the contrast facilitated learning, the 'yes-no' task would eliminate the possibility of comparing the suffixes as they heard the test items.

The third issue of the results across the experiments surrounds learning representations and phonology. As discussed in the previous section, §5.3, the rate of selecting the grammatical suffix significantly dropped in the MAT condition in both New Items and Old Items when there was a greater number of morphemes to learn, while there was no such decrease in the GP condition. Based on the current results, it is expected that people would have more difficulty learning in the MAT condition than in the GP condition when there are a higher number of morphemes to learn. This can be investigated and confirmed using the current experimental method by increasing the number of suffixes to learn. If this turns out to be the case, it would support the P-First model, especially when learning multiple morphemes at the same time. In other words, with more representations to learn, it is expected to be easier for the learner when a language has phonologically systematic and predictable alternations than when memorizing or listing all the representations with selection rules. This would provide behavioral evidence for the the P-First model.

There are other ways to study this topic as well. One could test how long it takes or how much exposure one needs when learning multiple morphemes and the phonology of the language. This could be studied through tracking the accuracy rate based on the exposure to training items (Chandrasekaran et al., 2014a,b; Linzen and Gallagher, 2014). The paradigm measures how fast participants reach the ceiling. Again, if the P-First model is correct, people will reach the ceiling faster when learning the GP grammar, then the results would provide behavioral evidence for GP.

If increasing the number of morphemes that participants have to learn results

in better learning of the GP-grammar, this would provide indirect evidence regarding learning the pronunciations of the morphemes in natural language. When learning a natural language, infants are not exposed to one morpheme at a time. Rather, they are exposed to multiple morphemes with multiple pronunciations at the same time. Thus, abstracting phonological generalizations with a single representation for each morpheme may be less demanding than listing all the representations for each morpheme with phonologically-conditioned selection rules. This would also answer the question of how much information or input we need to learn the morpheme alternant patterns and to abstract phonological generalizations.

Although there is no clear answer at this moment, this section has provided a discussion of remaining questions and directions for future studies.

Chapter 7

CONCLUSION

This dissertation investigated the relationship between the lexical representations of various pronunciations of morphemes and their learnability. The topic was studied by using an artificial grammar learning paradigm. The mental representations of allomorphs have been argued as having either a single underlying representation or multiple representations. In this dissertation, the former view was referred to as GP and the latter as MAT, following Kenstowicz and Kisseberth's (1979) terminology.

This dissertation has made the following contributions.

First, this dissertation shows that one can use an AGL paradigm to take underlying representations into consideration when studying the learnability of allomorphy patterns. Previous research using the artificial grammar learning paradigm to study phonology mainly focused on the learning of phonotactics in terms of phonetic naturalness, or formal complexity. Although previous studies provide important information about learning of phonology, the question asked in this dissertation is fundamental to phonology.

Second, this dissertation draws attention to experimental design. Across the experiments, the results were robust and replicable. The results, therefore, provide strong evidence in favor of studying the learning of multiple pronunciations of morphemes using a two-alternative forced-choice task. In other words, this dissertation shows that it is possible to study the nature of representation using behavioral experiments.

Third, this dissertation provides several directions for future experiments. Future studies should examine cognitive load during the process of learning the allomorph patterns. The difference between categorizing stem types in the MAT condition and

the GP condition is one of the interesting topics to be probed. Adding more suffixes will reflect how people learn or categorize phonologically analyzable and unanalyzable suffixes.

Most important are the results from the AGL experiments in this paper. The results from this series of experiments consistently reject the MAT-view and the M-First model, which predict that there will be no difference in learning. The results support the GP prediction that there still be a difference in learning. Furthermore, the results do not support the P-First model of morpho-phonological learning and instead point to a more complex interaction between phonological and morphological learning. In this way, this dissertation provides a foundation and direction for future research into the structure of mental representations of allomorphs in our minds.

BIBLIOGRAPHY

Ackerman, F., Malouf, R., and Blevins, J. P. (2016). Patterns and discriminability in language analysis. *Word Structure*, 9(2):132–155.

Allen, B. and Becker, M. (2015). Learning alternations from surface forms with sublexical phonology. Under revision for *Phonology*. http://ling.auf.net/lingbuzz/002503.

Anderson, S. (1985). *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Linguistics. University of Chicago Press, Chicago.

Basbøll, H. (2005). *The Phonology of Danish*. The Phonology of the World's Languages. Oxford University Press, New York, NY, USA.

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

Becker, M. and Gouskova, M. (2016). Source-Oriented Generalizations as Grammar Inference in Russian Vowel Deletion. *Linguistic Inquiry*, 47(3):391–425.

Berent, I. (2013). The phonological mind. *Trends in cognitive sciences*, 17(7):319–27.

Berko, J. (1958). The child's learning of english morphology. *Word*, 14:150–177.

Blevins, J. P., Ackerman, F., and Ackerman, F. (2016). Morphology as an adaptive discriminative system. In Siddiqi, D. and Harley, H., editors, *Morphological Metatheory*, pages 271–301. John Benjamins Publishing Company, Amsterdam; Philadelphia.

Chambers, K. E., Onishi, K. H., and Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, 87(2):B69–B77.

Chandrasekaran, B., Koslov, S. R., and Maddox, W. T. (2014a). Toward a dual-learning systems model of speech category learning. *Frontiers in Psychology*, 5(July):1–17.

Chandrasekaran, B., Yi, H. G., and Maddox, W. T. (2014b). Dual-learning systems during speech category learning. *Psychonomic Bulletin and Review*, 21(2):488–495.

Cho, S. W. (1982). The acquisition of word order in Korean. In *Calgary Working Papers in Linguistics*, pages 53–116.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. Harper & Row, New York.

Faust, N. and Lampitelli, N. (2016). Allomorphy—its logic and limitations: opening remarks from the guest editors. *Morphology*, 26(3-4):229–234.

Finley, S. (2012). Typological Asymmetries in Round Vowel Harmony: Support from Artificial Grammar Learning. *Language and cognitive processes*, 27(10):1550–1562.

Finley, S. (2015). Learning nonadjacent dependencies in phonology: Transparent vowels in vowel harmony. *Language*, 91(1):48–72.

Finley, S. (2017). Learning metathesis: Evidence for syllable structure constraints. *Journal of Memory and Language*, 92:142–157.

Finley, S. and Badecker, W. (2010). Linguistic and non-linguistic influences on learning biases for vowel harmony. In Ohlsson, S. and Catrambone, R., editors, *Proceedings of the 32nd Annual Conference of the cognitive science society*, pages 706–711, Austin, Texas. Cognitive Science Society.

Fitch, W. T., Hauser, M. D., and Chomsky, N. (2005). The evolution of the language faculty: clarifications and implications. *Cognition*, 97(2):179–210; discussion 211–25.

Folia, V., Udd, J., Forkstam, C., and Petersson, K. M. (2010). Artificial Language Learning in Adults and Children. *Language Learning*, 60(December):188–220.

Gainor, B., Lai, R., and Heinz, J. (2012). Computational characterizations of vowel harmony patterns and pathologies. In *The Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 63–71.

Gerken, L. A. and Knight, S. (2015). Infants generalize from just (the right) four words. *Cognition*, 143:187–192.

Gómez, R. and Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in cognitive sciences*, 4(5):178–186.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598):1569–1579.

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., and Mehler, J. (2011). Consonants and vowels: different roles in early language acquisition. *Developmental Science*, 14(6):1445–1458.

Hwangbo, H. J. (2015). An experimental study of the learnability of two vowel harmony patterns with neutral vowel. Unpublished Manuscript.

Hyman, L. (1975). *Phonology: Theory and Analysis*. Holt McDougal, New York.

Hyman, L. M. (1979). *Aghem grammatical structure*. Number No. 7 in Southern California Occasional Papers in Linguistics. University of Southern California, Los Angeles.

Hyman, L. M. (2011). Tone: Is it Different? In Goldsmith, J. A., Riggle, J., and Yu, A. C. L., editors, *The Handbook of Phonological Theory: Second Edition*, chapter 7, pages 197–239. Wiley-Blackwell, Malden, MA, 2nd edition.

Hyman, L. M. (2018). Why underlying representations? *Journal of Linguistics*, pages 1–20.

Inkelas, S. (2014). *The Interplay of Morphology and Phonology*. Oxford University Press, New York, NY, USA.

Kaun, A. R. (2004). The Typology of Rounding Harmony. In Hayes, B., Kirchner, R. M., and Steriade, D., editors, *Phonetically Based Phonology*, chapter 4, pages 87–116. Cambridge University Press, Cambridge ; New York.

Kenstowicz, M. and Kisseberth, C. (1979). *Generative Phonology*. Academic press, San Diego, California.

Kim, Y.-j. (1997). The Acquisition of Korean. In Slobin, D. I., editor, *The Crosslinguistic Study of Language Acquisition. Vol. 4*, chapter 4, pages 335–444. Lawrence Erlbaum Associates, Inc.

Kiparsky, P. (1982). *From Cyclic to Lexical Phonology*, volume I, pages 131–175. Foris, Dordrecht, Netherlands.

Lai, R. (2015). Learnable vs. Unlearnable Harmony Patterns. *Linguistic Inquiry*, 46(3):1–55.

Lai, Y. R. (2012). *Domain specificity in learning phonology*. Dissertation, University of Delaware.

Linzen, T. and Gallagher, G. (2014). The timecourse of generalization in phonotactic learning. *Proceedings of the Annual Meetings on Phonology*, pages 1–12.

Marcus, G. F. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, 283(5398):77–80.

Marslen-Wilson, W. and Tyler, L. K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences*, 2(11):428–435.

Marslen-Wilson, W. D. and Tyler, L. K. (1997). Dissociating types of mental computation. *Nature*, 387(June):592–592.

Marslen-Wilson, W. D. and Tyler, L. K. (2007). Morphology, language and the brain: the decompositional substrate for language comprehension. *Phil Trans R Soc Lond B Biol Sci*, 362(1481):823–36.

McClelland, J. L. and Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences*, 6(11):465 – 472.

McNaughton, R. and Papert, S. (1971). *Counter-free automata*. M.I.T. Press, Cambridge, Mass.

Moreton, E. and Pater, J. (2012a). Structure and Substance in Artificial-phonology Learning, Part I: Structure. *Language and Linguistics Compass*, 6(11):686–701.

Moreton, E. and Pater, J. (2012b). Structure and Substance in Artificial-Phonology Learning, Part II: Substance. *Language and Linguistics Compass*, 6(11):686–701.

Nevins, A. (2011). Phonologically conditioned allomorph selection. In van Oostendorp, M., Ewen, C. J., Hume, E., and Rice, K., editors, *The Blackwell Companion to Phonology*, volume IV of *Blackwell companions to linguistics series*, chapter 99, pages 2357 –2382. Wiley-Blackwell Publishing, West Sussex, United Kingdom.

No, G. (2009). Acquisition of case markers and grammatical functions. In Lee, C. and Simpson, G., editors, *The Handbook of East Asian Psycholinguistics*, chapter 3, pages 50–62. Cambridge University Press, New York.

Odden, D. (2014). *Introducing Phonology*. Cambridge University Press, New York, United States of America, 2nd edition.

Onishi, K. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, 83(1):B13–B23.

Öttl, B., Jäger, G., and Kaup, B. (2015). Does formal complexity reflect cognitive complexity? Investigating aspects of the Chomsky Hierarchy in an artificial language learning study. *PloS one*, 10(4):e0123059.

Peperkamp, S. and Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. *Laboratory Phonology*, 9:315–338.

Peperkamp, S., Skoruppa, K., and Dupoux, E. (2006). The Role of Phonetic Naturalness in Phonological Rule Acquisition. In Bamman, D., Magnitskaia, T., and Zaller, C., editors, *Proceedings of the 30th Annual Boston University Conference on Language Development (BUCLD 30)*, pages 464–475, Somerville, MA. Cascadilla Press.

Petersson, K. M., Forkstam, C., and Ingvar, M. (2004). Artificial syntactic violations activate broca's region. *Cognitive Science*, 28:383–407.

Pinker, S. (1998). Words and rules. *Lingua*, 106(1):219–242.

Pinker, S. and Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6(11):456–463.

Prince, A. and Smolensky, P. (1993). Optimality theory: Constraint interaction in generative grammar. RuCCS-TR-2. 262pp. ROA-537.

Psychology Software Tools, Inc. (2016). E-prime 2.0. http://www.pstnet.com.

Pycha, A., Nowak, P., Shin, E., and Shosted, R. (2003). Phonological Rule-Learning and its implications for a theory of vowel harmony. In Garding, G. and Tsujimura, M., editors, *Proceedings of the West Coast Conference on Formal Linguistics*, volume 22, pages 423–435, Sommerville, MA. Cascadilla.

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Ramscar, M., Dye, M., and McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language*, 89(4):760–793.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6):855–863.

Rogers, J., Heinz, J., Fero, M., Hurst, J., and Wibel, S. (2013). Cognitive and Sub-Regular Complexity. In *Formal Grammar*, pages 90–108.

Rogers, J. and Pullum, G. K. (2011). Aural Pattern Recognition Experiments and the Subregular Hierarchy. *Journal of Logic, Language and Information*, 20(3):329–342.

Scobbie, J., Coleman, J., and Bird, S. (1996). Key aspects of declarative phonology. *Current Trends in Phonology: Models and Methods*, 2:685–709.

Seidl, A. and Buckley, E. (2005). On the Learning of Arbitrary Phonological Rules. *Language Learning and Development*, 1(3-4):289–316.

Seidl, A., Cristià, A., Bernard, A., and Onishi, K. H. (2009). Allophonic and Phonemic Contrasts in Infants' Learning of Sound Patterns. *Language Learning and Development*, 5(3):191–202.

Selkirk, E. (1982). The syllable. In van der Hulst, H. and N., S., editors, *The structure of phonological representations (Part II)*, pages 337–383. Foris Publication.

Silva, S., Folia, V., Hagoort, P., and Petersson, K. M. (2017). The P600 in Implicit Artificial Grammar Learning. *Cognitive Science*, 41:137–157.

Sober, E. (2015). *Ockham's razors.* Cambridge University Press, Cambridge, United Kingdom.

Sohn, H.-M. (1999). *The Korean Language.* Cambridge University Press, Cambridge, United Kingdom.

Steriade, D. (2008). A pseudo-cyclic effects in Romanian morphonology. In Bachrach, A. and Nevins, A., editors, *Inflectional identity*, chapter 10, pages 313–359. Oxford University Press, Oxford.

Toro, J. M., Shukla, M., Nespor, M., and Endress, A. D. (2008). The quest for generalizations over consonants: Asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception and Psychophysics*, 70(8):1515–1525.

Trubetzkoy, N. (1969). *Principles of Phonology*. University of California Press, Berkeley.

Tyler, L. K., DeMorany-Davies, P., Anokhina, R., Longworth, C., Randall, B., and Marslen-wilson, W. D. (2002a). Dissociations in Processing Past Tense Morphology : Neuropathology and Behavioral Studies. *Journal of Cognitive Neuroscience*, 14(1):79–94.

Tyler, L. K., Randall, B., and Marslen-Wilson, W. D. (2002b). Phonology and neuropsychology of the English past tense. *Neuropsychologia*, 40(8):1154–1166.

Ullman, M. T. (2015). The Declarative / Procedural Model: A Neurobiological Model of Language. In Hickok, G. and Small, S. L., editors, *Neurobiology of Language*, chapter 76, pages 953–968. Elsevier Science.

White, J. and Sundara, M. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition*, 133(1):85–90.

Wilson, C. (2003). Experimental Investigation of Phonological Naturalness. In Garding, G. and Tsujimara, M., editors, *WCCFL*, volume 22, pages 533–546, Sommerville, MA.

Wilson, C. (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive science*, 30(5):945–82.

Yang, C. D. (2000). *Knowledge and Learning in Natural Language Certified by Accepted by*. PhD thesis, Massachusetts Institute of Technology.

Appendix A

TRAINING ITEMS

	Traning Stem GP		Training suffix MAT			Traning Stem GP		Training suffix MAT	
CVC	bos	boset	bos	bosut	rVC	ræb	ræbet	ræk	rækut
	gæm	gæmet	gæm	gæmut		ræk	ræket	ræs	ræsut
	kem	kemet	kem	kemut		ræs	ræset	reg	regut
	kut	kutet	kut	kutut		reg	reget	rem	remut
	nep	nepet	nep	neput		rem	remet	rin	rinut
	nik	niket	nik	nikut		rob	robet	rip	riput
	nub	nubet	nub	nubut		rod	rodet	rod	rodut
	pæg	pæget	pæg	pægut		rok	roket	rok	rokut
	sig	siget	sig	sigut		run	runet	run	runut
	tob	tobet	tob	tobut		rut	rutet	rut	rutut
CVr	ber	beræt	ber	beræt	rVr	rær	ræræt	rær	ræræt
	bur	buræt	bur	buræt		rer	reræt	rer	reræt
	gær	gæræt	gær	gæræt		ror	roræt	ror	roræt
	gur	guræt	gur	guræt		rur	ruræt	rur	ruræt
	kir	kiræt	kir	kiræt				rir	riræt
	mer	meræt	mer	meræt					
	pir	piræt	pir	piræt					
	por	poræt	por	poræt					
	sor	soræt	sor	soræt					
	tær	tæræt	tær	tæræt					

Appendix B

TEST ITEMS

New items

	GP	MAT		GP	MAT		GP	MAT
CVC	bep	bep	CVr	bor	bor	rVC	ræd	ræd
	bis	bis		dær	dær		ræg	ræn
	bok	bok		dur	dur		ræn	ræp
	dæt	dæt		ger	ger		ræp	reb
	deg	deg		gir	gir		reb	rek
	gub	gub		kær	kær		rek	ret
	kæm	kæm		kor	kor		rep	rig
	kib	kib		kur	kur		ret	rim
	mek	mek		mor	mor		rog	ris
	mut	mut		ner	ner		rom	rog
	pob	pob		nir	nir		ros	rom
	pon	pon		nur	nur		rot	rot
	sib	sib		per	per		rud	rud
	suk	suk		sær	sær		rup	rup
	tæs	tæs		tir	tir		rus	rus

Old items

	GP	MAT		GP	MAT
CVC	kem	kem	rVC	ræs	ræs
	nub	nub		rep	rep
	sig	sig		rob	rob
	tob	tob		rut	rut
		pæg			rip
CVr	bur	bur	rVr	rær	rær
	gær	gær		rer	rer
	mer	mer		ror	ror
	pir	pir		rur	rur
		sor			rir

Appendix C

FULL INSTRUCTIONS OF THE TRAINING SESSION AND TEST SESSION OF THE EXPERIMENTS

<Training Instructions>

You will hear a list of word pairs in a foreign language.

Each pair contains a word and then the same word with the plural suffix.

Please repeat each pair of words after they are presented.

Once you have repeated a pair, press SPACEBAR to proceed to the next

pair.

This will last approximately 20 minutes without any breaks,

And you will not be able to stop the training until it is finished.

When the training session is completed, you will enter the test session.

Press SPACEBAR to begin training.

Figure C.1: Instructions for training session

```
<Test instruction 1/2>


                You will hear several rounds of words.
      Each round includes three words:  an unsuffixed word followed by two
                          suffixed words.
           You are asked to judge whether the first suffixed word
    or the second suffixed word is more likely to belong to the language
                          you just heard.
              Press SPACEBAR to see more Test Instructions.
```

```
<Test instruction 2/2>


    If you think the first suffixed word you just heard is grammatical, press
                              ''1''.
       If you think the second suffixed word you just heard is grammatical,
                          press ''0''.
         Please position your left finger on '1' and right finger on '0'.
                   Press SPACEBAR to start the TEST.
```

Figure C.2: Instructions for test session

Appendix D

IRB APPROVAL

UNIVERSITY OF DELAWARE **RESEARCH OFFICE**

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: June 16, 2017

TO: Hyun Jin Hwangbo
FROM: University of Delaware IRB

STUDY TITLE: [830703-6] Phonology and MAT

SUBMISSION TYPE: Amendment/Modification

ACTION: APPROVED
APPROVAL DATE: June 16, 2016
EXPIRATION DATE: December 10, 2017
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(7)*

Thank you for your submission of Amendment/Modification materials for this research study. The University of Delaware IRB (HUMANS) has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that <u>informed consent</u> is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

Generated on IRBNet

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.