

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Phonemic Variability in Word Learning

Permalink

<https://escholarship.org/uc/item/5cn6b9z8>

Author

Frye, Conor

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Phonemic Variability in Word Learning

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Cognitive Science

by

Conor Frye

Committee in charge:

Professor Sarah Creel, Chair
Professor Seana Coulson
Professor Marc Garellek
Professor Tamar Gollan
Professor Marta Kutas

2018

The Dissertation of Conor Frye is approved, and it is acceptable in quality and form for
publication on microfilm and electronically:

Chair

University of California San Diego

2018

DEDICATION

There is not enough space to offer my full appreciation for the help I have received throughout this process here. Suffice it to say, without the herculean patience and support first and foremost of my wife, Elizabeth Parks, my mother and father, and my sister, this document would never have occurred. I also have to thank those in my lab and the graduate students in the Cognitive Science program (and outside of it) who spent countless hours with me, answering questions, providing insight, and accompanying me to basketball games and coffee runs. The networks of interaction and support that naturally occurred throughout this process were indispensable.

Thank you all very much, I owe you a great debt, and you forever have my gratitude.

EPIGRAPH

I think about more than I forget,
but I don't go around fire expecting not to sweat-

Dwayne Michael Carter, Jr.

TABLE OF CONTENTS

Signature Page.....	iii
Dedication.....	iv
Epigraph.....	v
Table of Contents.....	vi
List of Figures.....	vii
List of Tables.....	ix
Vita.....	x
Abstract of the Dissertation	xi
Introduction	1
Chapter 1: Phoneme Variability in Word Learning.....	18
Chapter 2: Native Language and Phoneme Variability in Word Learning.....	33
Chapter 3: Differential Phonemic Attributes Effect.....	50
Chapter 4: Phonemic Variability: a More Sensitive Measure.....	65
Chapter 5: Developmental Effects on Phonemic Merging.....	76
Chapter 6: Conclusion.....	90
Bibliography	98

LIST OF FIGURES

Figure 1. Accuracy in each merge condition, with by-subjects standard errors. Dissimilar label condition was different from every other condition, collapsed across language background, $p < .001$. Chance performance = .50 23

Figure 2. Collapsed across language backgrounds, close trials were significantly different from all other conditions ($p < .001$), same consonant trials were significantly different from same vowel trials ($p < .001$). The dissimilar-label condition is not included26

Figure 3. Pearson’s product-moment correlation of overall test accuracy and Spanish participants’ score on the Spanish Bilingual Dominance Scale (BDS). There was no significant correlation between score and accuracy ($p = .47$) nor was it strongly directional ($r = -.12$).....29

Figure 4. Accuracy in each merge condition, separated by language background. Collapsed across backgrounds, the dissimilar-label condition was different from the onset-merge condition ($p < .001$), the onset control condition ($p < .001$), the vowel-merge condition ($p = .01$) and the vowel control condition ($p = .002$)..... 40

Figure 5. Accuracy on test based on trial type split by language background. When collapsed over background, close trials were significantly different from all same onset consonant trials ($p < .001$), which were significantly different from same vowel trials ($p < .001$), which were significantly different from far trials ($p = .03$). The dissimilar-label condition is not included.....41

Figure 6. Pearson’s product-moment correlation of overall test accuracy and participants’ score on the Spanish Bilingual Dominance Scale (BDS). There was no significant correlation between Spanish BDS score and accuracy ($p = .42$) nor was it strongly directional ($r = -.13$).....43

Figure 7. Proportion of correct responses to post-test trials split by language background. When collapsed, accuracy was significantly better for different phoneme trials than different consonant trials ($p < .001$) as well as for different consonant trials versus different vowel trials ($p = .002$).... 45

Figure 8. Accuracy in each merge condition. Dissimilar label condition was different from every other condition: Onset Merge ($p = .02$), Coda Merge ($p = .004$), Onset Control ($p < .001$), and Coda Control ($p < .001$)..... 55

Figure 9. Close trials were significantly different from same onset trials ($p < .001$), which were significantly different from same vowel trials ($p < .001$). The dissimilar-label condition is not included here..... 57

Figure 10. Proportion of correct responses to post-test trials was significantly better for different phoneme trials than different consonant trials ($p < .001$) as well as for different vowel trials ($p < .001$)..... 62

Figure 11. Accuracy in each merge condition. Dissimilar label condition was different from every other condition, $p < .001$ 68

Figure 12. Close trials were significantly different from all same consonant trials ($p < .001$), which were significantly different from same vowel trials ($p < .001$), which were significantly different from far trials ($p = .01$). The dissimilar-label condition is not included here..... 69

Figure 13. Accuracy in each merge condition. Collapsed across phoneme type, merging conditions were different from Single-label Control conditions, $p = .01$ 81

Figure 14. Correct-other proportion of looks in each merge condition. Collapsed across phonemes, merging labels received a smaller proportion of correct looks than Single-label Control conditions, $p < .001$ 83

Figure 15. Pointing accuracy in each condition. Children performed above chance regardless of the minimal pair type, $p = .002$ 86

Figure 16. Correct-Other proportion of looks in each minimal pair condition. Neither condition showed a higher proportion of correct looks to target than chance, $p = .73$ 87

LIST OF TABLES

Table 1. A complete list of the nonce words used in the experiment.....	20
Table 2. A complete list of the nonce words used in the experiment.....	37
Table 3. A complete list of the nonce words used in the experiment.....	52
Table 4. A collection of results of each experiment in this work and their insight into the flexibility of word learners in specific flexible phoneme situations based on the comparison of merging, segment, and trial type effects.....	95

VITA

2011 Bachelor of Arts, University of Southern California
2011-2012 Teaching Assistant, University of California San Diego
2013 Master of Science, University of California San Diego
2013-2017 Research Assistant, University of California San Diego
2018 Doctor of Philosophy, University of California San Diego

FIELDS OF STUDY

Major Field: Cognitive Science

Studies in Language Acquisition
Professor Sarah Creel

ABSTRACT OF THE DISSERTATION

Phonemic Variability in Word Learning

by

Conor Frye

Doctor of Philosophy in Cognitive Science

University of California San Diego, 2018

Professor Sarah Creel, Chair

Language learning processes are often examined by learning miniature languages in the lab, where controlling the learning of a full language is infeasible. Most work on learning words assumes these meaningful chunks of sound are composed of smaller, equally meaningful phonemes. However, recent evidence suggests crystallized phoneme categories do not fully explain how words are learned but instead to flexible representation of phonemes in terms of gradient similarity in word learning by adults and children.

I propose a shift in the understanding of the role of phoneme categories in word learning by presenting a set of three studies consisting of 6 experiments that challenge our understanding of phoneme representation. The first study assesses whether the adult word learning system is more flexible in its word representations than previously described and finds repeated examples that within certain boundaries such flexibility is easily achieved. The second study relies on eye-

tracking measures, a more sensitive measure of recognition difficulty, to extend the behavioral results from the first and demonstrate that recognition is similarly rapid for phonemically-consistent and phonemically-variable words. It illuminates parts of the underlying responses of the adult word-learning system. The third study explores how flexibility changes across developmental time, testing 3-5-year-old children on their ability to learn flexible vocabularies as well as distinguish minimal pairs. This final study suggests that though children are capable of learning phonemically-variable words, the flexibility experienced by adults is due to extended time, practice, or experience with a language, and not available immediately to learners.

These findings contribute to the general field of word learning across three main avenues: first, by demonstrating incredible flexibility in sound category interpretation across adult learners of different linguistic backgrounds, counter to prevailing assumptions, and adding nuance to the research on bilingual advantages in word learning. Second, contributing new accuracy and time-course data on positional and segmental importance to speech sound flexibility in word learning, adding to the general understanding of word learning. And finally, by providing data to extrapolate a developmental time course for these abilities, contributing to our picture of spoken language development over a developmental time scale.

INTRODUCTION

Learning language is one of the most complex and intriguing feats that humans achieve; but how we do it remains obscured behind many layers of questions we have yet to answer. Given its complexity, many researchers explore facets of the language-learning puzzle by using simplified tasks in the laboratory. One aspect of this problem that has been explored substantially using laboratory experiments as a microcosm is word learning. In principle, by exploring how different populations learn a small number of words under highly controlled conditions, we can build up to an understanding of language acquisition in the real world.

In their approach to this problem hitherto, many researchers have regarded words as being composed of discrete sound-based symbols, or *phonemes*. If each word is a unique combination of these specific, unchanging building blocks, then word learning can be understood as a process of learning these building blocks, and then combining them to form words. In this account, a word is learned when a new sequence of symbols is stored (and becomes associated with a meaning). Using this theory, the study of word learning has arrived at specific conclusions for how word learning takes place, but often does not seem capable of explaining all the nuances of actual word learning and speech perception.

This document proposes a shift in the current understanding of the role of phoneme categories in word learning and more generally in language learning. By demonstrating that the conclusions reached in previous research are only part of the story, this document will challenge the current understanding of word learning and bring more nuanced understanding to the field. In a set of five experiments, I aim to assess whether the adult word learning system is more flexible in its representations than previously described; whether phoneme recognition—once understood

to be a critical determinant of word identity—is much more adaptive than previously thought; and how such flexibility changes across developmental time.

The phoneme as the atom of word form identity

In the field of language acquisition, there is a major assumption that a learner’s initial goal is to detect specific sound categories in that language. In English, for example, the sounds that differentiate the word “cat” from the word “cut” are considered different phonemes. That is, exchanging the “a” for the “uh” sound is a clue to word learners that they are hearing a new word. These symbol-like sound categories, known as phonemes, indicate differences in meaning between words (Liberman et al., 1957). Once discrete phonemes are learned, speakers can efficiently code learned words and identify when they are hearing a new word.

Many models of language conceptualize the phoneme as elemental to word identity. Best’s perceptual assimilation model (PAM; Best, 1994; Best, McRoberts, & Goddell, 2001) suggests that second language phoneme perception is driven by articulatory similarity to the learner’s first-language phonemes. If a phoneme in a second language is different from the native language, perceptual accuracy will diminish. This approach strongly argues for the phoneme as a discrete symbolic unit of word identity. Throughout the literature, there is an assumption that phoneme category boundaries are fairly fixed in adulthood (though see Flege, 1998, for a lifelong-learning perspective). These studies and theoretical accounts imply that learners should have difficulty learning phonemically-identical words as different—for instance, “ba” and nasalized “ba,” which—while phonemically different in French—are not phonemically different in English. This appears to be true (Silbert, Smith, Jackson, Campbell, Hughes, & Tare, 2015; Pajak, Creel, and Levy, 2016). However, it is worth noting that words with *similar* phonemes are

also more difficult to learn as distinct words (e.g. Creel et al., 2006, 2008; Creel & Dahan, 2010; Magnuson, Dixon, Tanenhaus, & Aslin, 2007), suggesting that learning difficulty may increase with degree of sound similarity even apart from phoneme identity. This in itself raises doubts as to the utility of phoneme boundaries.

Just as these accounts imply that learning distinct words cannot occur when they differ only on a non-phonemic attribute, they also imply that learners should have difficulty learning phonemically-*differing* words as the *same* word—for example, learning that paff and baff are equally-acceptable labels for a novel concept. Since these are two different sequences of phonemes, they should need to be learned as separate words. This second implication has been much less explored than the first one, and is the subject under consideration in this proposal.

Evidence for the rigidity of phoneme categories

As the current understanding goes, when novices learn specific phonemes in a language, they are acquiring the tools to recognize and encode known and new words in that language. The language learner's most basic and important task is then to learn the categories of separate phonemes—find the phonemes, find the words. Thus, a great deal of research has focused on the perceptual problem of distinguishing between highly similar phonemes (Delattre, 1967; Abramson, Lisker, 1970; Lisker, 1978; Flege, 1984; Flege and Hillenbrand, 1984).

Perhaps the paradigm case of phoneme category learning is the stop consonant voicing distinction. For example, the sounds “b” as in “beach” and “p” as in “peach” are highly similar, differing mainly in subtle timing parameters. One of these parameters, called “voice onset time” or VOT, is a major contributor to the category distinction. It is this difference that defines these two sounds, and creates a boundary between them. Listeners are aware of this boundary as well

as the acoustic information included in the speech stream. Reaction times were faster when responding to tokens with similar acoustic information on the same side of the border, as well as when distinguishing tokens with large acoustic differences on different sides of the VOT boundary (Pisoni & Tash, 1974). This boundary may be learned through exposure to distributions of VOTs in the speech signal, and the field has argued for developmental changes in categorical boundaries as well (e.g. Hazan & Barrett, 2000; though see Eimas et al., 1974, for an account of developmentally-early sensitivity to the VOT boundary). With a creation of a boundary, two distinct categories are created for use in language learning.

By some accounts, native-language phoneme categories become entrenched early in life. Although most languages have sound categories roughly analogous to “b” and “p,” VOT cues (and related cues) differ from language to language, meaning that cues and their associated boundaries must be learned specifically for one’s native language. Infants are thought to be born with the neural capacity to discriminate nearly all phonemes in a language-universal manner (Werker & Tees, 1984), but during their first twelve months, linguistic exposure causes a pruning of perceptual abilities so that only native language phoneme categories remain (Kuhl et al., 2006). The older the learner is when a language is learned, the poorer the learner’s ability to perceive and produce distinctions between these categories (Flege, 1991). This means that non-native speakers may often confuse (for example) a non-native language’s “b” sound with a “p” sound, or vice versa, because of the inflexibility of comparison with their native categorical boundary perception.

Further, native-language phoneme categories are difficult to alter when one is learning a new language. In one particularly well-studied case, native Japanese speakers find it notoriously difficult to distinguish an English “r” sound from an English “l” sound (Miyawaki et al., 1975).

This fits with the symbol-like category structure in that it is consistent with adult learners having rigid learned boundaries for their native contrasts, and later contrast testing demonstrates an inability to modify these native symbol boundaries without substantial training. Different languages even have different numbers of categories. Spanish, for example, is considered to have five vowels (Maddieson, 1984), whereas standard American English has been noted to have 11 (Clopper et al., 2005). For native bilingual speakers and those learning a second language, this difference in category number might make moving between languages difficult, and could cause problems in new word learning if the phoneme categories do not perfectly match up. For example, if a native Spanish speaker wants to learn an English word with a vowel that falls between two known Spanish vowels, what is the best course of action? Is it easier to treat the new vowel category as an instance (albeit a poor one) of similar known categories, or to learn an entirely new set of boundaries for a new category? In the symbolic categorical phoneme account, these questions remain unanswered.

Evidence for flexibility of phoneme categories

Other evidence suggests that phonemic representation is not so rigidly defined as the above literature implies. When learners are purposefully given multiple labels for a single object (e.g. “div” and “dev”) in an experimental setting, they learn as readily as they learn a single label (“dev”), suggesting highly-flexible phoneme boundaries, or at least graded sensitivity to similarity across phoneme boundaries (Muench & Creel, 2013). This means that when learning labels that contain phoneme variability, in certain circumstances adults have little more difficulty than with learning a single label. But why people can learn multiple labels as easily as they can

learn one in certain circumstances is not entirely understood. And notably, in accounts where phonemes are symbolically distinct from their neighbors, this result is difficult to account for.

There is similar evidence for flexibility even in the most prototypical example of categorical speech perception, voice onset time (e.g. McMurray, Tanenhaus, Aslin, 2002). McMurray et al. (2002) asked participants to find a specific picture (e.g. a peach) whose spoken label varied in exact VOT. Phonological competitors (e.g. a beach) were placed in the visual field that differed only across a category boundary (peach/beach). Participants were more likely to categorize long-VOT “b” sounds as short-VOT “p” sounds and vice versa, indicating that the categorical boundary which seemed so stark for perception of a single phoneme is actually much less categorical when given natural words (McMurray et al., 2002). Further, even when a listener chose (say) the peach, if its VOT was closer to the boundary, they showed more visual fixations to the beach than when VOT was further from the p/b boundary. It is clear that categorical perception should be possible here for word distinctions, but it may not extend to more realistic comprehension. The flexibility displayed in these examples lends evidence to the argument that understanding word learning through a categorical or symbolic lens is incomplete.

Segment type and phoneme category flexibility

A relevant consideration here is that phoneme category flexibility may vary by segment type (consonants vs. vowels). Consonants outnumber vowels in most languages, are normally shorter in time, are less continuous, and often differ in voicing, formant structure, and perceptual quality (Owren & Cardillo, 2005). Consonants and vowels differ from each other in how categorically they are perceived, and the surrounding sound context that a phoneme occurs in modifies its categorical nature to different extents. For instance, in a study asking participants to

compare a continuum of stimuli to a standard sound, both consonants and vowels showed flexibility based on which type of standard was presented, but vowels demonstrated a much larger range of acceptable alteration (Macmillan, Goldberg, & Braida, 1988). This is evidence that there might be additional aspects of word learning affecting how these category boundaries are perceived, and importantly suggests that vowels, perhaps more so than consonants, might be more amenable to that sort of instability.

Consonants and vowels have been described phonetically as two separate categories (Maddieson, 1984; Ladefoged, 2001). This separation has led to research demonstrating that different languages have different roles for consonants vs. vowels: most languages have more consonants than vowels (Maddieson, 1984). Some researchers have gone so far as to argue that consonants determine word meaning while vowels are more critical for syntactic categories (Nespor, Peña, & Mehler, 2003), implying that vowels are not as useful for word learning. Of course, Nespor et al.'s account assumes that all languages have the same relationship between vowels and consonants.

Crucially, vowels and consonants have been demonstrated cross-linguistically to provide speakers with different information about new words. Cutler, Sebastian-Galles, Soler-Vilageliu, and Van Ooijen (2000) found that learners look to consonants for word recognition more reliably than vowels. Cutler et al. gave participants non-words such as “teeble” and asked them to turn them into real words by changing single sounds, in this example, “table” or “feeble.” When participants were told to change the vowels of the non-words, accuracy was higher and response speed was faster; given the choice, participants more often altered vowels than consonants. This was true in both Spanish and Dutch (languages with low and high vowel to consonant ratios, respectively). This difference in informational content was further explored in a statistical

learning task given to Italian adults. Alternating whether information about syntactic rules or lexicon information were given by vowels or consonants, Toro, Nespor, Mehler, and Bonatti (2008) concluded that word segmentation is driven by consonants, whereas the syntactical information was carried in the vowels.

Studies on word learning have demonstrated a tendency to ignore vowel differences between words, which one might interpret as greater vowel flexibility. For instance, English-speaking adults learning novel words were much less likely to confuse words distinguished by consonants (such as *pibo* and *zigo*) than to confuse words distinguished by vowels (such as *pibo* and *pabu*; Creel, Aslin, & Tanenhaus, 2006). Yet Creel et al. (2006) also found evidence that words differing in their vowels were more distinct than those differing in their *coda* consonants. This implies that the greater “lexicality” or inflexibility of consonants may depend on syllable position, with onset consonants more critical to word identity than coda consonants.

This pattern continues in lexical priming tasks. When primed by spoken non-words, French speakers were globally faster to respond to a target word when it contained the same consonant as the prime word (Delle Luche et al., 2014). Delle Luche et al. also found that English speakers, given stress-pattern priming as well as vowel and consonant patterns, showed similar results, save for vowel priming demonstrated in iambic words. This means that English speakers alone were affected by priming with stress patterns as well. When stress was placed on the last syllable, something uncommon in English, vowels provided more information to the lexicality decision. To explain this result, they argue that rhyme preservation explains this new attention to vowels in English speakers, and that the stress priming highlighted the rhyme. These studies demonstrate that, on the whole, word information is preserved more when vowels are altered. Although they did not directly test word learning, this finding argues that vowel

variability should be less confusing when learning new words. However, it also does not rule out flexibility within consonants in certain structures or in certain positions within words.

Developmental changes in phoneme categories

Another aspect of this puzzle is how phoneme-category flexibility may change over developmental time. Behavioral studies with children as well as adults suggest that phonemic categories are not set in stone by late infancy, specifically, that phoneme changes do not always signify word-meaning changes (for some examples of flexibility in children, see Swingley and Aslin, 2000, 2002; Poltrock & Nazzi, 2015; Havy & Nazzi, 2009; Nazzi & Bertoncini, 2009; Nazzi & New, 2007; for some examples of flexibility in adults, see McMurray, Tanenhaus, & Aslin, 2002; Creel, Aslin, & Tanenhaus, 2006; Muench & Creel, 2013). This should not be possible if word learning occurs by checking for known and unknown phoneme sequences in order to detect new words. Learners' apparent allowance for phonemic "wobble room" even after infancy and into adulthood may point to a neural architecture for language acquisition that operates in graded probabilities (degree of similarity) rather than symbol-like phoneme representations. This would mean that instead of phonemes being symbol-like elements with rigid category boundaries, they are malleable probability spaces that can adapt, even after the native language is learned.

This flexibility has been noted in developmental populations as well as adults. Substantial evidence has been presented by Nazzi and colleagues to demonstrate the flexibility of vowels in native perception. In a set of related studies, French infants from 11 to 20 months recognized words after vowel changes (Poltrock & Nazzi, 2015; Havy & Nazzi, 2009; Nazzi & Bertoncini, 2009; Nazzi & New, 2007). For each of those studies, Nazzi and colleagues relied on a specific

task they developed where infants were presented with a trio of dissimilar objects by an experimenter, two with the same name, and one differing in different manners depending on the study. Infants were then asked to categorize the objects based on the label given by the experimenter. Infants tested in those studies tended to prefer known words, and rely more on the consonant information remaining prototypical than the vowels. However, this may not be the case for all age groups, as 5-month-old infants seemed to rely more on *vowels* in their name remaining constant, allowing the consonants to shift without detriment to its understanding (Bouchon, Floccia, Fux, Adda-Decker, Nazzi, 2014). The consonant-bias, therefore, would seem to emerge late in the first year of life.

Related work examining children's familiar word recognition (Swingley & Aslin, 2000, 2002) presents young children with either correct pronunciations ("doggy") or incorrect pronunciations that deviate by one phoneme ("toggy"). When hearing "toggy," infants are more likely to look to a picture of a "doggy" (Swingley & Aslin, 2000, 2002) than to a picture of a shoe, though looks are most accurate when they hear a correct pronunciation ("doggy"). While those authors did not interpret their findings this way, one might infer that English infants as young as 14 months are sensitive to the similarity between a mispronunciation, even a consonant mispronunciation, and its base word.

Developmental flexibility in phonemes appears to continue past infancy. Preschool aged children demonstrate this phonemic flexibility as well: when presented with minimally different non-words ('fesh' instead of 'fish'), 3- to 5-year olds overwhelmingly pointed to and looked at a picture of a fish, rather than pointing to an unfamiliar object (Creel, 2012). This suggests that, to young children, 'fesh' is nearly identical to 'fish,' even though the two forms differ by a categorically different phoneme ("ih" vs. "eh"). A later experiment by Creel (2012) suggested

that preschool-aged children were just as flexible in recognizing familiar consonant-mispronounced as vowel-mispronounced words. Again, we see that speech sounds are somewhat flexibly interpreted well after infancy.

However, preschool children's phonemic flexibility may be somewhat restricted to familiar words. This same age group was tested on their ability to generalize *novel* (newly-learned) words from one artificial accent to another, as well as their ability to simultaneously learn two different versions of the same novel word (Creel, 2014). Children only sometimes showed generalization when the novel words underwent a minimal vowel change (/i/ to /I/ or the reverse) (Creel, 2014). When provided multiple speakers for the different accents learning novel words, children again seemed to learn talker-specific pronunciations instead of generalizing, implying they were paying more strict attention to each phoneme in the novel words, thus slowing down learning. It also suggests that children—unlike adults in Muench and Creel (2013)—may find it difficult to learn multiple similar forms of words, as accuracy in general declined from an earlier experiment where each word had a single pronunciation. Children in later experiments performed much better at recognizing accented forms of familiar words containing /i/ and /I/. Interspersing familiar words also mildly benefited their accuracy when learning novel words, implying some reliance on previously known vowel relationships when learning novel words in multiple accents (Creel, 2014).

These results are consistent with a theory proposed in Sebastián-Gallés et al.'s (2005, 2009) work that states that when learning multiple accents, learners (adults in their case) encode all word-forms separately. For example, when trained that the words 'teev' and 'tihv' are the same object, learners would separately encode both versions instead of learning to allow multiple imperfect versions of a single prototype. If this is what children are doing, it suggests that they

are not very *flexible* in phoneme perception, but that they need an increased *memory capacity* for phonemically related words.

Across a range of developmental studies, early language learners show flexibility for at least some phonemes in the right scenarios, implying a more generalized ability to recognize words' sound patterns in a flexible manner. Some studies suggest greater vowel than consonant flexibility (Nazzi and colleagues), while other studies imply flexibility for both types of sounds (Creel, 2012; Swingley & Aslin, 2000, 2002). This flexibility across age groups further demonstrates that greater elasticity in word acquisition is possible than would be predicted by categorical symbolic accounts.

Phonemic Flexibility in Multilingual Speakers

According to a 2006 European Commission survey, the majority of the world (~56%) speaks more than one language, making it of substantial interest to know how bilingualism affects phoneme flexibility. There is no question that being bilingual presents the language learning and perception system with certain challenges. There has been ample study on bilingualism and its beneficial effect on executive control (see Bialystok, Craik, Green, & Gollan, 2009 for a review). Exploring how bilinguals might be affected during word learning, however, has been less studied. Early bilingualism has been associated with facilitated phonological processing, possibly leading to better word learning (Bialystok et al., 2003). This improved phonological ability, however, has been linked to degree of relatedness between languages (see Bialystok et al., 2003; Bialystok et al., 2005). Could phoneme flexibility be affected by phonological processing facilitation? Might facilitation of phoneme flexibility, like facilitation of phonological processing, depend on the relationship between known languages?

When it comes to actual word learning, a second language may have beneficial effects. For instance, adult bilinguals performed better than adults who had never learned another language when given a task that paired a novel word with a known word (Papagno & Vallar, 1995; Van Hell & Mahn, 1997). In theory, this advantage might result from knowing the speech sound inventories of two languages, giving bilingual learners a larger set of sounds to use to encode word forms. However, one study found that bilinguals were at an advantage in word learning tasks that taught words with sounds novel to all languages known by the participants (Kaushanskaya & Marian, 2009). How this advantage will affect learning under phoneme variability is unknown. On the one hand, if bilingual speakers' previously-observed word learning benefit was due to increased ability to pair specific sound patterns with concepts, phoneme variability could lower their ability to learn new words that do not sustain a constant phoneme pattern. On the other hand, if bilingual speakers' previously-observed word learning benefit is driven by better acquisition of new sounds, as suggested by Kaushanskaya & Marian (2009), then bilinguals might find phoneme flexibility easier to learn than their monolingual counterparts.

Differentiating between these two outcomes will further our understanding of how word learning happens across varied language backgrounds. Knowing how these different backgrounds affect learning specifically will provide insight into how phonemes are learned generally as well as how new words are perceived by learners of different linguistic backgrounds.

Open Questions

To sum up, current research leaves questions about the extent to which phonemes are the atoms of word form identity; to what extent phoneme flexibility in word learning depends on the type and syllable position of the phoneme; whether knowing multiple languages changes these relationships; and the developmental trajectory of phoneme flexibility in word learning.

We know there is some flexibility in phonemes for word form identity, but the extent of that flexibility is as of yet unclear. It remains to be seen if all types of phonemes are flexible to the same degree, or if flexibility is limited to vowels alone. Research on consonants and vowels leaves open two questions with respect to phoneme category flexibility. The first is the extent to which specific types of phonemes (vowels or consonants) are flexible within word recognition. Secondly, we are currently unclear on the positional effect on each type of phoneme. It seems that for consonants, syllable onsets may be relatively inflexible, but whether this inflexibility extends to coda consonants is unclear.

It is clear that different languages use relationships between phonemes differently when giving meaning to word forms. What is as of yet not known, and what we ask here, is if this difference occurs in a measurable manner based on basic properties of the language, such as the number of vowels and consonants in a particular language.

Finally, although some studies have explored developmental changes in phoneme flexibility, more information is needed as to the specific trajectory of flexibility during development, particularly for learning new words. It is unclear if children are less flexible during word learning than adults, and if that is the case, why this would be so.

Experiment Summary

With these questions in mind, this document presents two complementary lines of studies on word learning. The first set of studies explores adults' phoneme flexibility in word learning within and across languages, with implications for adult plasticity during language learning and more general word learning capability (and phonemic flexibility) in adults. Three related experiments investigate this question. The first experiment asks monolingual and bilingual groups of participants to learn English-based vocabularies with either onset consonant variability or vowel variability. This study explores the question of how crucial stability in phonemes is for learning words. The second experiment in this series explores differences in language background further by presenting subjects with a vocabulary designed to resemble Spanish. This study provides insight into the generality of phoneme flexibility, and demonstrates to what extent the previous results are due to language-specific processes vs. more general language learning processes. The final study in this trio probes the roles of phoneme type (consonant, vowel) and position (onset, coda) in learning flexibility by comparing variable-vowel learning to variable-coda learning. This experiment sheds light on differences in amount of flexibility for different phonemes by exploring each interaction of phoneme type and position on the accuracy of learning a novel vocabulary.

The second goal of this document is to explore how phonemic malleability might lead to subtle processing difficulties that broader measures of accuracy miss and how those internal changes alter language learning and perception. The second line of study investigates this question in three experiments. The first explores implicit recognition of phonemically variable words through visual-world eye tracking. This study delves into the processes at work during word learning, and provides insight into how the language learning system treats flexibility in

phonemes while learning a new vocabulary. The second experiment relies on three to five-year-old learners to examine the role of development in the same processes during word learning explored in the first experiment. It explores how capable developing language learners are of learning from phonemically-variable input, increasing our understanding of language learning in its most basic sense from development through maturity in regard to linguistic flexibility. This study records behavioral and eye-tracking data from preschoolers performing a simplified version of adult tasks in earlier experiments, allowing more comparability between age groups than across previous studies (e.g. Muench & Creel, 2013, vs. Creel, 2014). It provides greater insight into the developing language learner's flexibility with respect to phoneme boundaries and the possible underlying mental structures that bring that skill to bear. A third experiment, also relying on three to five-year-old learners, examines the baseline capability of that age group to distinguish minimal pairs in order to rule out an alternative interpretation of the second experiment.

Potential Impact

A view of phonemes as the necessary and sufficient building blocks of word representation is simplistic and outdated. These studies will allow us to better understand the role of phonemes during word learning, and how phoneme flexibility contributes to learning language. This information has been hinted at by current literature, but has never been robustly, directly, and holistically explored. This document provides a solid empirical foundation for understanding the role of the phoneme in word learning across development, language background, phoneme type, and syllable location. Results provide useful evidence for multiple

fields of study, including first and second language acquisition, language development, and bilingual vocabulary acquisition.

Chapter 1: Phoneme Type Variability in English-based Word Learning

How crucial are phonemes for differentiating words? Experience and historical precedent recognize the phoneme as a critical determinant of word identity during language acquisition: *dog* is a word, but *tog* is not. Most language acquisition models describe phonemes as crystallizing into fixed categories by early childhood (Kuhl, 2000), making word learning as simple as identifying new words by their component speech sounds. However, other evidence suggests that phoneme representations are more flexible. Children can recognize words even if a constituent phoneme is violated (“cat” pronounced as “ket”; Creel, 2012), and adults can learn labels for the same object differing in vowel phonemes (e.g. *ziv/zev*) as readily as they learn a single label (Muench & Creel, 2013).

The first experiment tests the hypothesis that categorical use of phonemes is not rigidly defined in adult brains and behavior but can be modified with minimal training. Our *flexible perceptual probability hypothesis* predicts that category use is much more plastic and can be modified and merged. It predicts that exposure to different probability distributions should alter phoneme category use behaviorally. The current experiment seeks to overturn the primacy of the phoneme in favor of a more probabilistic definition of word identity, by teaching participants novel vocabularies with various “collapsed” phonemic differences—for example, teaching learners that a particular referent can be labeled both “deev” and “teev”—in order to gauge the malleability of phoneme representation at a behavioral level.

Method

Participants

We tested 80 normal-hearing adults recruited from the UC San Diego SONA subject pool. Forty were native English speakers (avg. age: 21.1 years, sd: 1.9) who grew up hearing no other languages and reporting fluency in no other language but English. All subjects were chosen based on a lack of exposure to any language other than American English before age 14, and a lack of any exposure or formal training in any language for more than 3 years after that point. The remaining 40 subjects were bilingual Spanish-English speakers (avg. age: 21.1, sd=2.9) who grew up hearing Spanish from birth, and English either at birth or sometime later. Age of English acquisition ranged from 0-14 years, and the average age was 5.0 years, sd=3.0. All bilingual speakers considered themselves fluent in both languages. Each subject participated in a prescreen asking specific questions about their language background and history with exposure to linguistic stimuli. This prescreen was used to calculate a bilingual dominance score (BDS) in the manner presented in Dunn and Fox Tree (2009). Spanish speaking participants averaged a score of 15.85 (SD= 5.51) on the Spanish dominance scale, and 20.93 (SD=3.64) on the English dominance scale, falling within Dunn and Fox Tree's narrow limit of ± 5 points for balanced bilingualism

In addition to the 80 subjects used in this data set, 26 additional subjects were run, but were dropped: four due to a coding error in one condition that was discovered during the experiment, seven due to excessive exposure to a second language, one due to experimenter error, and 14 due to subjects failing to follow task instructions.

Stimuli

Visual stimuli were a set of 16 two-dimensional monochromatic nonsense shapes first used by Creel et al. (2006). Auditory stimuli were 32 novel consonant-vowel-consonant (CVC) words. They were recorded in a soundproof recording room by one male and one female native English speaker in the Center for Research and Language at UCSD. During word-learning training and test phases, only recordings from the male speaker were used. During the post-test, recordings from both speakers were used. The vocabulary (Table 1) was designed in 4-word sets. These sets allowed construction of word pairs with either onset consonants that differed by voicing (“zuf” and “suf” were one pair), or that differed by the tenseness of the vowel (“zuf” and “zuhf” were one pair).

Table 1. A complete list of the nonce words used in the experiment

fosh (fʌʃ)	vosh (vʌʃ)	bish (bɪʃ)	pish (pɪʃ)
fush (fʌʃ)	vush (vʌʃ)	beesh (biʃ)	peesh (piʃ)
fehʃ (fɛʃ)	vehʃ (vɛʃ)	buhv (bʊv)	puhv (pʊv)
fayf (feɪf)	vayf (veɪf)	boov (buv)	poov (puv)
zadz (zʌdʒ)	sadz (sʌdʒ)	dehdz (dɛdʒ)	tehdz (tɛdʒ)
zudz (zʌdʒ)	sudz (sʌdʒ)	daydz (deɪdʒ)	taydz (teɪdʒ)
zoof (zuf)	soof (suf)	dihv (dɪv)	tihv (tɪv)
zuhf (zʊf)	suhf (sʊf)	deev (di:v)	teev (ti:v)

Design

There were five between-subjects conditions. In the *vowel merge* condition, labels for specific objects varied in the tenseness of their vowel phonemes (e.g. the same object was labeled both “zuf” and “zuhf”). In the *onset merge* condition, the onset consonant was variable (e.g. “zuf” and “suf”)—varying in exactly the voicing contrast that is the paradigm case of phoneme categorization. In the *dissimilar* (difficult) condition, all phonemes except the coda varied between the two labels for a single object (e.g. “zuf” and “fehʃ”). There were two

additional control conditions, where each picture received only one label. The *vowel control* conditions were matched to the vowel-merge condition by using either all tense vowels (only /u, i, eɪ, a/ used; half of participants) or all lax vowels (only /ʌ, ʊ, ɪ, ε/ used; half of participants) present in the vocabulary. Like the vowel merge condition, then, there were only four functional vowel categories, but unlike the vowel merge condition, phonemes were consistent. The onset consonant control was matched similarly to the onset consonant-merge version in that the control lists used only voiced onset consonants (/b,d,v,z/) or unvoiced onset consonants (/p,t,f,s/) to create the 16-word set. These control conditions guard against the possibility that having less sound category diversity in the set of to-be-learned words makes words more confusable.

Procedure

The study consisted of vocabulary training with feedback, followed by vocabulary testing (no feedback). In training, participants learned labels for 16 novel objects over 256 trials in a two-alternative forced choice task (2AFC). On each trial, participants saw two novel objects on a screen, and heard a recorded word. They chose one of the objects by clicking on it, and received accuracy feedback before advancing to the next learning trial. Specifically, the correct response picture remained onscreen while the incorrect response disappeared. An additional mouse-click advanced the learner to the next training trial. Non-target pictures had phonetically dissimilar labels to target pictures during training trials.

Next, each participant was tested (128 trials) on the trained word-meaning pairings in the same 2AFC task as during training, except that there was no feedback. Accuracy and reaction times were measured. Although training trials only depicted dissimilar-sounding words for competing objects on a given trial, the test trials contained four different levels of sound

similarity between target and competitor. In *close-competitor* trials, the target and the competitor shared two of the three phonemes in the word; in the onset consonant-merge and onset consonant-control conditions, target and competitor shared codas and onsets and differed in vowels, e.g., target: /fɑʃ/-vɑʃ/; competitor: /fʌʃ/-vʌʃ/. For the vowel-merge and vowel-control conditions, target and competitor shared coda and vowel and differed in onsets, e.g., target: /fɑʃ/-/fʌʃ/; competitor: /vɑʃ/-vʌʃ/. There were also *same onset* and *same vowel* conditions, where target and competitor shared only an onset or vowel, respectively (target: /fɑʃ/-vɑʃ/ competitor: /fɛʃ/-vɛʃ/; target: /fɑʃ/-vɑʃ/; competitor: /zɑʒ/-sɑʒ/). Finally, in the *far* competitor test trials, either no sounds were shared or only the coda was shared (target: /fɑʃ/-vɑʃ/; competitor: /bʊv/-/pʊv/ or target: /fɑʃ/-vɑʃ/; competitor: /bɪʃ/-pɪʃ/).

Results

We created generalized mixed-effects models with the package lme4 (Bates et al. 2014, R Core Team 2014). Throughout this study, when any two models are compared, finding that one accounts for significantly greater variance in a likelihood test (anova() function) between fixed effects describes the significance of each fixed effect. This comparison is achieved by creating two models: one model is made with all fixed and random effects present, and another identical model is created with the exception of a single fixed effect. The chi-square statistic and its associated p-value are reported for this comparison. Furthermore, Barr et al. (2013) suggest that likelihood ratio tests provide a more reliable estimate of statistical significance than Wald's z-test.

The first model assessed whether there was a general effect of label similarity. That is, do we find—consistent with previous studies (Muench & Creel, 2013)—that learning dual labels is

easier when the dual labels are phonologically similar to each other than when they are not? To test this, we modeled test accuracy in terms of the fixed effects of condition type (onset control, onset merge, vowel control, vowel merge, and dissimilar labels) and background (Spanish-English bilingual or English monolingual). Fixed effects were sum-coded so that fixed effects patterns could be interpreted as they are in ANOVAs. Individual subject intercepts and word intercepts and slopes (both fixed effects and their interaction) were included as random effects. The dissimilar-labels learning condition was set as the reference level for this analysis. There was a main effect of condition type ($\beta=.62$, $SE=.17$, $Z=3.58$, $p<.001$) (Figure 1).

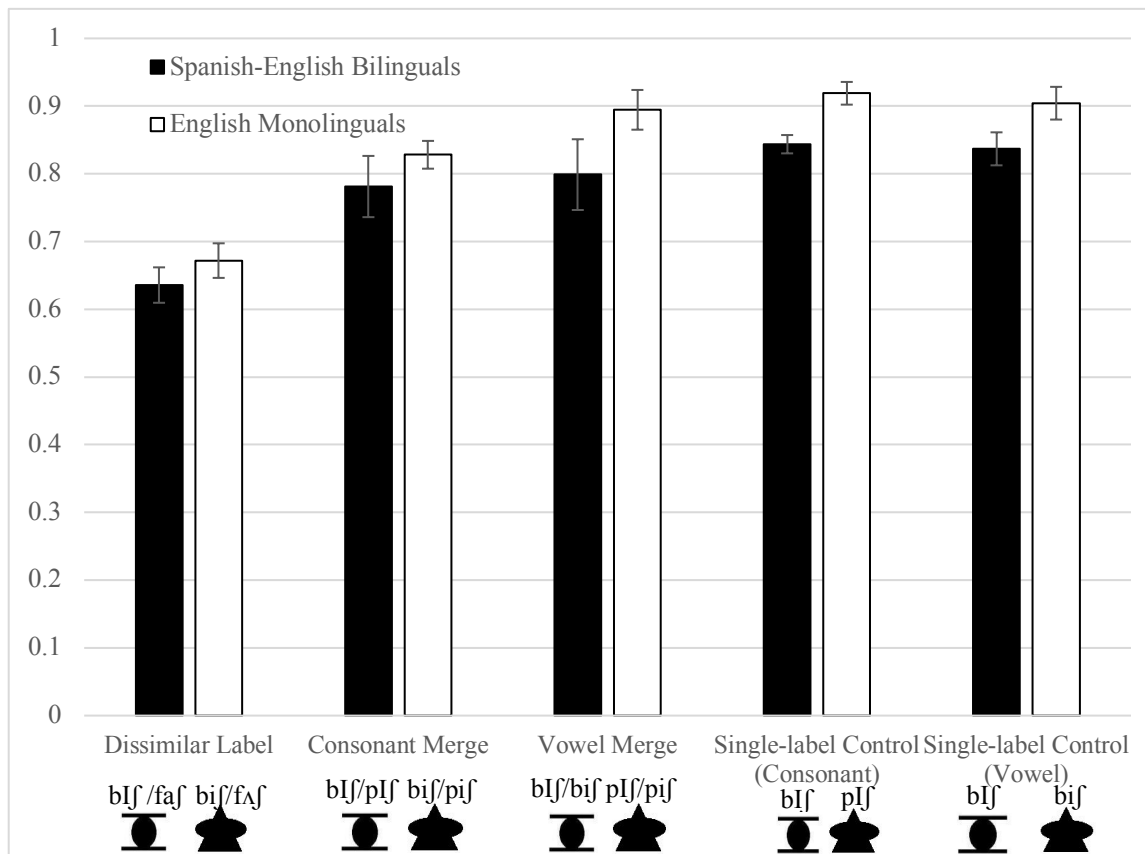


Figure 1. Accuracy in each merge condition, with by-subjects standard errors. Dissimilar label condition was different from every other condition, collapsed across language background, $p<.001$. Chance performance = .50.

This effect was driven by the dissimilar-labels condition showing lower accuracy than other conditions. Individually, each condition was significantly more accurate than the dissimilar labels condition ($p < .001$ for all comparisons between the dissimilar condition and any other condition).

The second model asked more specific questions about the difficulty of learning with variable labels. The dissimilar labels condition was dropped for this model. The effects tested were participant language background, merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and trial type (either sharing no phonemes, sharing a coda and an onset, sharing a coda and a vowel, or sharing onset and vowel), with participant accuracy as the dependent variable. Trial type was forward difference coded, so the contrasts compare trials with successively less-phonologically-similar competitors: close trials to same onset consonant trials, same onset consonant trials to same vowel trials, and same vowel trials to far trials (no shared phonemes). There was a main effect of background ($\beta = 0.33$, $SE = .09$, $Z = 3.79$, $p < .001$), and the model that removed the background condition effect was significantly different from the larger model ($\chi^2 = 13.70$, $p = .001$), with English monolinguals outperforming Spanish bilinguals. However, there were no interactions between background and other effects.

We then built a final model that removed background and compared merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and trial type (either sharing no phonemes, sharing a coda and an onset, sharing a coda and a vowel, or sharing onset and vowel) on participant accuracy. Again, trial

type was forward difference coded, so the contrasts compare trials with successively less-phonologically-similar competitors. The dissimilar labels condition was dropped for this model as well. There was a main effect of merge type ($\beta=0.28$, $SE=.13$, $Z=2.16$, $p=.03$), implying participants learning merged vocabularies were less accurate than those learning control vocabularies. The model that removed the merge effect ($\chi^2=5.51$, $p=.02$) was significantly different from the larger model, implying merging vocabularies were more difficult than controls. There was a main effect of trial type: removing the trial type effect was significantly different from the larger model ($\chi^2=82.02$, $p<.001$), implying that removing the trial type effects removed significant information from the larger model (Figure 2).

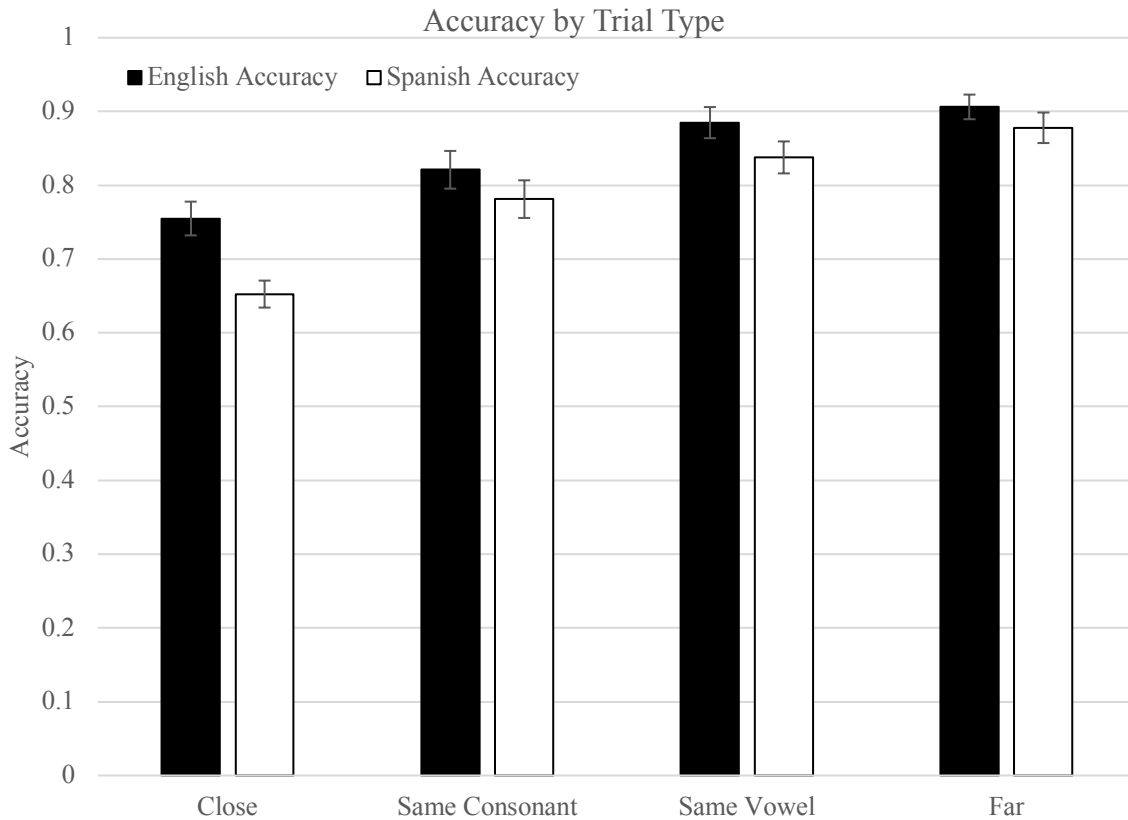


Figure 2. Collapsed across language backgrounds, close trials were significantly different from all other conditions ($p < .001$), same consonant trials were significantly different from same vowel trials ($p < .001$). The dissimilar-label condition is not included.

Specifically, participants performed less accurately when presented with competitors with each successive addition of phoneme sharing, with the most confusing being sharing 2/3 of their segments in close trials. More simply, performance on close trials was less accurate than same onset trials ($\beta = -1.15$, $SE = .17$, $Z = -6.98$, $p < .001$), which was subsequently less accurate than same vowel competitors ($\beta = -.66$, $SE = .17$, $Z = -3.84$, $p < .001$).

There were also two two-way interactions of segment condition and same onset and same vowel trial types and of merge condition and same onset trial type, though these are both qualified by a three-way interaction of same vowel trial type, segment, and merge type conditions ($\beta = .31$, $SE = .14$, $Z = 2.17$, $p = .03$). To discover what drove this interaction, we broke it

down into two simple interaction analyses at each level of segment type. First, we explored the interaction of merge type and trial type for onsets, and second that same interaction for vowels.

For onset conditions (onset merge and onset control), there was a main effect of merge condition ($\beta=.45$, $SE=.17$, $Z=2.62$, $p=.009$) with merging leading to lower accuracy, and trial types, where close competitor trials were less accurate than same onset trials ($\beta=-1.41$, $SE=.15$, $Z=-9.46$, $p<.001$), which were subsequently less accurate than same vowel competitor trials ($\beta=-1.33$, $SE=.30$, $Z=-4.47$, $p<.001$). There was also an interaction of same onset trial type and merge conditions ($\beta=-.29$, $SE=.14$, $Z=-2.07$, $p=.04$) with merging leading to lower accuracy during close competitor trials than during same onset competitor trials.

For vowel conditions (vowel merge and vowel control), there was a main effect of trial type, with accuracy on close competitor trials being lower than that on same onset competitor trials ($\beta=-.83$, $SE=.30$, $Z=-2.77$, $p=.006$), and accuracy on same vowel competitor trials being lower than that of far competitor trials ($\beta=-.92$, $SE=.30$, $Z=-3.07$, $p=.002$). However, there was no effect of merge type and no interaction. This suggests that, as in Muench and Creel (2013), learning words with inconsistent but highly-similar vowel phonemes does not impair learning accuracy.

We were also interested in behavior over the learning trials to determine if participants were differing in their trajectories of phoneme merging but eventually arriving at the same accuracy in test. The dissimilar labels condition was dropped for this model. The effects tested were merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and whether their training accuracy differed across the first half of the training session in comparison to the second half. We found a main

effect of training half ($\beta=.54$, $SE=.05$, $Z=11.67$, $p<.001$) and merge condition ($\beta=.32$, $SE=.10$, $Z=3.03$, $p=.002$). There were also two two-way interactions of merge condition and section half and of merge condition and segment type, though these are both qualified by a three-way interaction of section half, segment, and merge type conditions ($\beta=-.10$, $SE=.05$, $Z=-2.14$, $p=.03$).

To discover what drove this interaction, we broke it down into two simple interaction analyses at each level of section half. First, we explored the interaction of merge type and segment type for the first half, and second that same interaction for the second half of training trials. For first half training trials there was a main effect of merge condition ($\beta=.22$, $SE=.07$, $Z=2.94$, $p=.003$) with merging leading to lower accuracy. For second half training trials there was a main effect of merge condition ($\beta=.42$, $SE=.16$, $Z=2.53$, $p=.01$) as well as an interaction with merge condition and segment type ($\beta=-.35$, $SE=.16$, $Z=-2.20$, $p=.03$).

We also correlated overall accuracy with the Spanish BDS scores calculated earlier. Using Pearson's product-moment correlation, the BDS score did not predict a significant correlation between the accuracy data and the Spanish scores on the BDS ($t(38)=-.72$, $p=.47$). The relationship was not strongly directional ($r=-.12$). This relationship is demonstrated in Figure 3. The same correlation was made removing the participants in the dissimilar condition, and again, the BDS score did not predict a significant correlation between the accuracy data and the Spanish scores on the BDS ($t(30)=-.49$, $p=.63$). The relationship was not strongly directional ($r=-.09$).

Correlation of Overall Accuracy and Spanish BDS Score

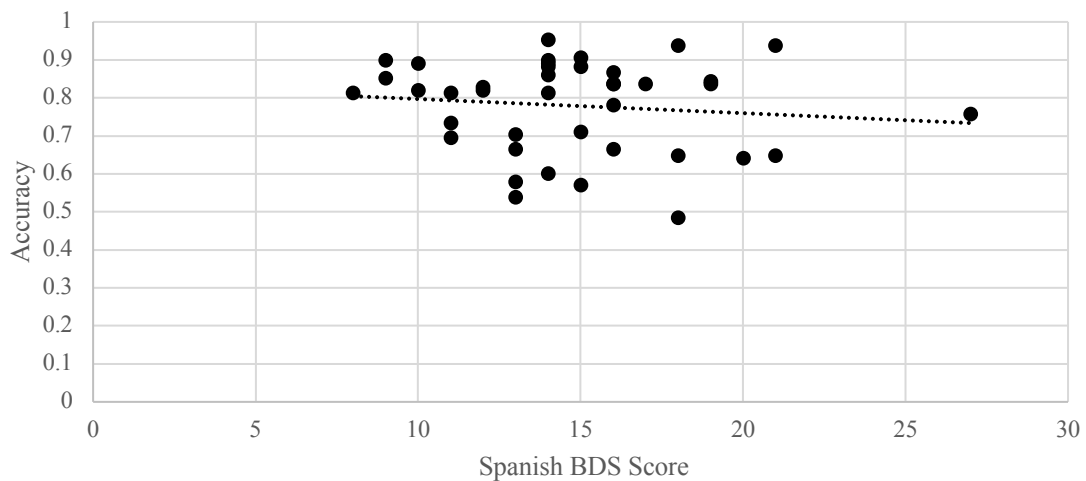


Figure 3. Pearson’s product-moment correlation of overall test accuracy and Spanish participants’ score on the Spanish Bilingual Dominance Scale (BDS). There was no significant correlation between score and accuracy ($p=.47$) nor was it strongly directional ($r=-.12$).

Discussion

Our findings contribute to research on word learning in two respects. Previous research showed that adults are capable of learning phonemically-variable vocabularies (Muench & Creel, 2013). Based on that work, we expected the *dissimilar labels* condition to be the most difficult, and the *single-label control* conditions to be the easiest, which was the case. The question was where the vowel merge and onset merge conditions would fall. Although there were not clear differences between onset consonant variability and vowel variability, the findings here suggest that learning words with variable onsets or variable vowels, while slightly more difficult than learning non-varying words, imposes limited difficulty relative to learning dissimilar labels. This implies there is a gradient of phoneme flexibility in word recognition and learning, with single-feature variation adding only mild difficulty to the learning process.

We also hypothesized that there might be greater learning accuracy among bilingual speakers, who are accustomed to switching between two phonological systems and thus might do so more easily in learning novel vocabulary. Many researchers have recently argued for cognitive benefits of bilingualism (Kaushanskaya & Marian, 2009; Bialystok, Craik, & Luk, 2009). For instance, previous research indicated that knowing a second language might convey some benefits for word learning (Kaushanskaya & Marian, 2009). Were this the case, then bilingual speakers should have performed better than monolinguals in our word-learning task. However, Spanish-English bilinguals uniformly performed *less* accurately. It is not immediately clear why our findings contradict earlier ones. This could be explained by cognitive factors unrelated to cross-linguistic phonological processing, or, as Muench and Creel (2013) suggested, Spanish-English bilinguals have lower familiarity with English phonology—the phonology of the to-be-learned words. Yet a third possibility is that Spanish-English bilinguals have access to more focused examples of good vowels than strictly English speakers, limiting the natural overlap in language and providing a lifetime of training to ignore borderline vowels. Because Spanish has fewer vowel categories than English, but necessarily fills the same vowel space, examples of vowels straddling borders between valid tokens due to speaker characteristics or muscular imprecision are likely fewer, providing less reason for a native Spanish speaker to pay attention to “in-between” vowels. This relative lack of vowels would also increase the likelihood that a misspoken vowel would lead to a meaningfully different word as opposed to a novel word that is recognized by the speaker as misspoken. Therefore, Spanish speakers might be less likely to merge vowel categories because it is simultaneously less likely to happen in that language and more damaging if a vowel is to fall out of place. The necessary information to specifically prove which potential hypothesis might be correct lies outside the scope of the current experiment.

In the current experiment, we predicted that merged vocabularies would be slightly more difficult to learn than control vocabularies, and this turned out to be the case. As the vocabulary in this study is more variable than what learners often encounter naturally, the outcome that systematic variability slightly lowers accuracy makes sense. However, as the merged vocabularies were still more accurate than vocabularies with dissimilar word pairs as object labels, it demonstrates a surprising flexibility in adult language learners. If phonemes truly were stable building blocks of language learning, participants given objects with labels that systematically vary across known language boundaries would be expected to face substantial difficulty at learning. This is not the case, implying that for language learning, even adult learners have a previously unreported level of flexibility in phoneme categorization when learning words.

We had also hypothesized that merging across variable phonemes would be more difficult when consonants varied than when vowels varied. However, differences between the segment types only differed significantly for particular test trial types. Critically, there was not a two-way interaction of merge type (merged or consistent sounds) and segment type (consonants or vowels) that would have indicated a general disadvantage for merging one segment type. Previous work has argued strongly for a consonant bias in English words, or at least an increase in mutability for vowels, but this experiment does not provide support for that argument (Cutler, Sebastian-Galles, Soler-Vilageliu, and Van Ooijen, 2000; Creel, Aslin, & Tanenhaus, 2006). This experiment found no difference in accuracy for participants merging consonants versus merging vowels as a whole, and only recorded a significant change between the segment types when challenging each with specific competitors. This outcome supports the idea that flexibility

in phonemes in word learning applies in a wide range of scenarios, and might only trigger significant differences given very specific and challenging tasks.

These results first add nuance to the ongoing debate regarding bilingual advantages in word learning: not every learning task benefits from knowledge of multiple languages and their contrasting sound categories, and in fact some interactions between languages might make learning more difficult. Second, they suggest impressive flexibility in adult word learning: not only variability in vowels—generally agreed to be flexibly perceived—but also, to some extent, variability in consonants differing categorically in voicing, can be learned. This hints at greater plasticity in L2 (and perhaps L1) acquisition than previously thought.

Chapter 2: Linguistic Fluency in Phonemic Processing

Does knowing the subphonemic information that is specific to a language provide improved performance to native speakers when learning new words? Recent studies have suggested that, in addition to cognitive advantages seen in bilinguals (see Bialystok, Craik, Green, & Gollan, 2009 for a review), there may also be a bilingual advantage in learning new languages. In fact, there is evidence that bilingual speakers have a specific advantage when learning words (Kaushanskaya & Marian, 2009). However, there is work that does not find this advantage, and in fact demonstrates the opposite (Muench & Creel, 2013, and the current Study 1 [Chapter 1]), with bilingual speakers learning slightly *less* well than monolingual speakers. To explain this result, one could argue that being bilingual places speakers at a disadvantage for processing the sound patterns of a less-dominant language (Muench & Creel, 2013). If this is the case, a vocabulary derived from the sound patterns of a specific language would give speakers of that language an advantage; creating a vocabulary from a different phonetic/phonological inventory should change how participant groups respond. Specifically, if the same language groups are used as in Chapter 1, a vocabulary created using Spanish phonemes and recorded by a native Spanish speaker should benefit Spanish-speaking participants more than monolingual English-speaking participants.

We predict that if native language phonological knowledge facilitates sound category merging, then native Spanish speakers (bilingual in English) should now show learning advantages over monolingual English speakers. However, if native Spanish speakers do not see a marked improvement over English speakers for a novel vocabulary based in Spanish phonemic space, then phoneme flexibility is caused by some other factor. These possibilities include the size of the phonemic space of the dominant language or the presence or absence of the phonemes

being modified in the original set. If specific native knowledge of a language improves word learning in flexible phoneme situations, it can be assumed that phoneme flexibility is improved by exposure to a range of examples of that phoneme within a language. This would demonstrate evidence for a much more focused adult phoneme flexibility to being within a native language, but would still demonstrate the plasticity of the adult language learning apparatus as it pertains to sound perception.

There is a second possibility as well. The results from the first and second experiments regarding bilingual performance might have to do with socio-economic status differences likely present in our populations, judging by what we know about the populations around the University of California, San Diego. Specifically, there is evidence Spanish-English bilinguals in southern California tend to be of more modest socioeconomic status than monolingual English speakers (Harrell & Carasquillo, 2003). Higher SES is correlated with richer linguistic input across development (Hoff, 2003), so it could be that monolinguals might outscore bilinguals in experiment one due to their enriched environment and language exposure. The current experiment can only demonstrate the importance of the phonemic space the novel vocabulary is constructed in to the learner. However, in conjunction with other studies in this document, it adds important evidence to the question of phonemic flexibility in word learning.

Method

Participants

We tested 80 normal-hearing adults recruited from the UC San Diego SONA subject pool. Forty participants were native English speakers (avg. age: 20.5 years, sd: 2.1) who grew up hearing no other languages and reporting fluency in no other language but English. These

monolingual participants had less than 3 years of high school language instruction in any other language besides English, and no participant reported exposure to any language besides English before age 14. The other 40 were bilingual Spanish-English speakers (avg. age: 20.1, $sd=1.5$) who grew up hearing Spanish from birth, and English either at birth or sometime later. Age of English acquisition ranged from 0-14 years, and the average age was 4.2 years, $sd=2.7$. All bilingual speakers considered themselves fluent in both languages.

Each subject participated in a prescreen asking specific questions about their language background and history with exposure to linguistic stimuli. Specifically, we measured when each participant began speaking a second language, and the amount of time they spent using that language. In the current study, the average age of acquisition of English as a second language was 4.23 years ($sd= 2.66$), with years being spent bilingual being, on average, 15.95 years ($sd= 2.56$). This prescreen was used to calculate a bilingual dominance score (BDS) in the manner presented in Dunn and Fox Tree (2009). Spanish speaking participants averaged a score of 18.05 ($SD= 5.11$) on the Spanish dominance scale, and 17.87 ($SD=3.78$) on the English dominance scale, well within Dunn and Fox Tree's limit of ± 5 points for balanced bilingualism.

Finally, due to an error in the lab, only 10 subjects were administered a MINT test (Gollan et al., 2012). Of the 10 given, six were given to participants with a Spanish-English bilingual background. For all participants, the test was administered in English to verify baseline fluency in English. If any participants mentioned exposure to Spanish, the MINT was also delivered in Spanish after the English version. Out of a possible 68 correct, English monolingual participants scored a mean of 62.50 on the English MINT, with a standard deviation of 0.58. Monolingual participants mentioning Spanish class exposure (3 participants) scored a mean of 2.50 on the Spanish MINT, with a standard deviation of 1.63. Spanish-English bilingual participants, all of

whom had heard Spanish from birth and English simultaneously or at some later point, scored a mean of 58.50 on the English MINT, with a standard deviation of 1.05, and a mean of 36 on the Spanish MINT, with a standard deviation of 13.19. Administration of the MINT test was more stringent in this study than in Gollan et al.'s original work, with only correct participant recall meriting scoring. No hints were given to any participants at any time, also differing from Gollan et al.'s original work. However, the BDS has been shown to be correlated with the MINT test in terms of performance on word-learning tasks (Quam & Creel, 2017), making it likely further MINT scores would corroborate the BDS scores and their correlation with accuracy.

In addition to the 80 subjects whose data is used in this study, 13 subjects were replaced, one due to the subject not completing the task correctly, four due to one Onset Change list possibly being harder because of sharing a second consonant, unlike other trials, and eight due to a coding error which displayed the same picture for two objects during some training sets in the Onset Control condition.

Stimuli

Visual stimuli were the same as Experiment 1. Auditory stimuli were 32 novel consonant-vowel-consonant-vowel (CVCV) words. These words were designed using Spanish phonemes and following Spanish phonotactic rules. The ending *u* was used so that all words were plausibly of the same gender (*u* is a rare but attested Spanish word ending that applies to some feminine and some masculine nouns (e.g. *el espíritu/ the spirit; la tribu/ the tribe*). They were recorded in a soundproof recording room by one male and one female native Spanish speaker in the Center for Research and Language at UCSD. The vocabulary was designed so that 16-pair sets could be made for the merge conditions. These sets were of word pairs with either

onset consonants that differed by voicing (/bivu/ and /pivu/ were one pair), or vowels close in Spanish vowel space (/fihu/ and /fehu/ were one pair). A complete list of the words used is presented in Table 2.

Table 2. A complete list of the nonce words used in the experiment

bevu (bevu)	belu (belu)	fehu (fexu)	fesu (fesu)
beevu (bivu)	beelu (bilu)	feeju (fixu)	feesu (fisu)
pevu (pevu)	pelu (pelu)	vehu (vexu)	vesu (vesu)
peevu (pivu)	peelu (pilu)	veehu (vixu)	veesu (visu)
davu (davv)	dalv (dalu)	gavu (gaxu)	gasu (gasu)
dovu (dovu)	dolu (dolu)	gohu (goxu)	gosu (gosu)
tavu (tavv)	talv (talu)	kavu (kaxu)	kasu (kasu)
tovu (tovv)	tolv (tolu)	kohu (koxu)	kosu (kosu)

Procedure

The design and procedure were the same as Experiment 1, except the learned vocabulary was different. In the *vowel merge* condition, labels for specific objects varied in their vowel phonemes (e.g. the same object was labeled both “pivu” and “pevu”). In the *onset merge* condition, the onset consonant was variable (e.g. “pivu” and “bivu”)—varying in exactly the voicing contrast that is the paradigm case of phoneme categorization (though note that in Spanish, voicing occurs as a pre-voiced vs. short-VOT-lag contrast). In the *dissimilar* (difficult) condition, all phonemes except the final vowel varied between the two labels for a single object (e.g. “bevu” and “gaju”). In the control conditions each picture received only one label. The

vowel control conditions were matched to the vowel-merge condition by using either both front vowels (e and i) or both back vowels (a and o) present in the vocabulary. The onset consonant control was matched similarly to the onset consonant-merge version in that the control lists used only voiced or unvoiced onset consonants to create the 16-word set. During training, competitor objects had names that were phonologically dissimilar to the target objects' names.

As in Study 1, there were four types of test trials. In close-competitor trials, the target and the competitor share three of the four phonemes in the word. For the consonant-merge and consonant-control conditions, target and competitor shared onsets and differed in first vowels, e.g., target: /**belu**/-/**pelu**/ competitor: /**bilu**/-/**pilu**/, whereas for the vowel-merge and vowel-control they shared vowels and differed in onsets, e.g., target: /**belu**/-/**bilu**/ competitor: /**pelu**/-/**pilu**/. In same-vowel and same-consonant trials, the target and competitor shared only the vowels or the onset consonant and final vowel, respectively (target: /**belu**/-/**bilu**/ competitor: /**fexu**/-/**fixu**/; target: /**belu**/-/**pelu**/ competitor: /**bivu**/-/**pivu**/). Finally, in the far trials, only the final vowel is shared (target: /**belu**/-/**pelu**/ competitor: /**visu**/- /**fisu**/).

Following the experiment proper, participants were given a 256-trial post-test where they were asked to make similarity judgments on two-word pairs. This was to assess whether participants were still perceptually distinguishing between similar sounds made by the speaker in the test portion of the experiment. Two native Spanish speakers, a male and the same female from the training and testing portion, recorded the original 32-word vocabulary. Each trial consisted of two spoken words, one from each speaker. Listeners had to judge whether the spoken words were the same word or different words. Using two different speakers on each trial made it impossible for listeners to do an exact acoustic match. Each word in the 32-word vocabulary was played in a word pair an equal number of times, counterbalanced for

presentation in the first and second position as well as speaker gender. Participants could respond by typing the letter “z” on the keyboard if the words they heard were the same, and the “m” on the keyboard if they were different. For each word, there were two trials where the exact same word was said by both speakers (once in each of the two possible speaker orders), and six trials where the words were different, in both possible speaker orders. In the different trials, the words could share all sounds but the second consonant, all sounds but the onset, or no sounds at all.

Results

The first model assessed whether there was a general effect of label similarity. To test this, as in Experiment 1, we modeled test accuracy in terms of the fixed effects of condition type (onset control, onset merge, vowel control, vowel merge, and hard). Individual subject intercepts and word intercepts and slopes (both fixed effects and their interaction) were included as random effects. The dissimilar-labels learning condition was set as the reference level for this analysis. There was an effect of condition with the dissimilar-labels being different from the onset-merge condition ($\beta=.79$, $SE=.24$, $Z=3.34$, $p<.001$), the onset control condition ($\beta=.79$, $SE=.24$, $Z=3.32$, $p<.001$), the vowel-merge condition ($\beta=.60$, $SE=.23$, $Z=2.56$, $p=.01$) and the vowel control condition ($\beta=.76$, $SE=.24$, $Z=3.17$, $p=.002$) (Figure 4).

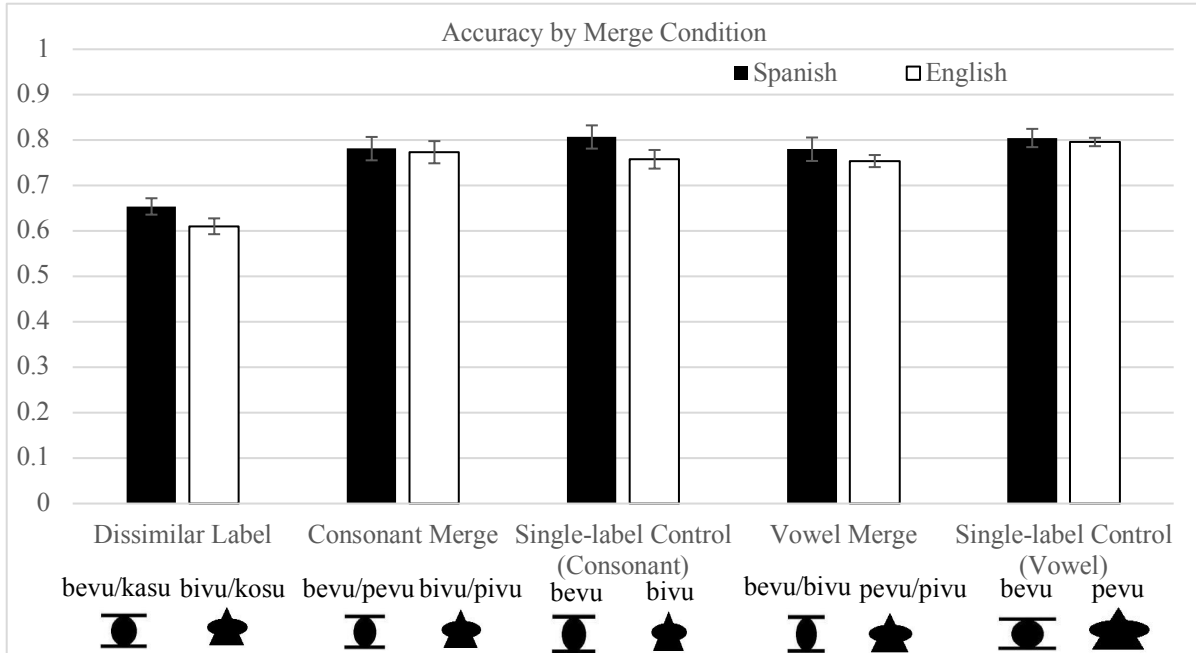


Figure 4. Accuracy in each merge condition, separated by language background. Collapsed across backgrounds, the dissimilar-label condition was different from the onset-merge condition ($p < .001$), the onset control condition ($p < .001$), the vowel-merge condition ($p = .01$) and the vowel control condition ($p = .002$).

We then explored a model which took into account subject language background, merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and trial type (either sharing no phonemes, sharing both consonants, sharing a second consonant and a vowel, or sharing onset and final vowel) on participant accuracy. That is, do we find that learning dual labels is easier when the dual labels are based in the phonology of the native language of the learner than when they are not? Trial type was forward difference coded, so the contrasts compare trials with successively less-phonologically-similar competitors: close trials to same onset consonant trials, same onset consonant trials to same vowel trials, and same vowel trials to far trials (no shared phonemes). The dissimilar labels condition was dropped for this model. There was a main effect of trial type

where close competitor trials were significantly less accurate than same onset competitor trials ($\beta=-.35$, $SE=.07$, $Z=-4.86$, $p<.001$), same onset competitor trials were significantly less accurate than same vowel competitor trials ($\beta=-.94$, $SE=.08$, $Z=-11.15$, $p<.001$), and those were also significantly less accurate than far competitor trials ($\beta=-.21$, $SE=.10$, $Z=-2.18$, $p=.03$). See Figure 5 for details. Overall, the model without trial type was significantly different than the larger model ($\chi^2=461.77$, $p<.001$).

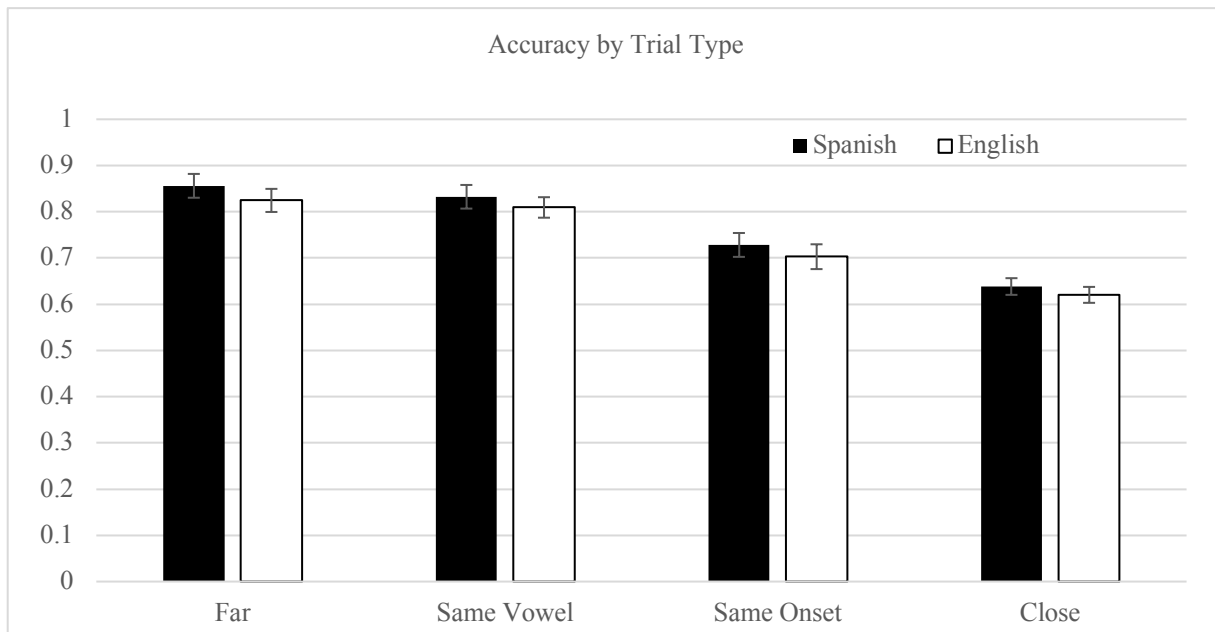


Figure 5. Accuracy on test based on trial type split by language background. When collapsed over background, close trials were significantly different from all same onset consonant trials ($p<.001$), which were significantly different from same vowel trials ($p<.001$), which were significantly different from far trials ($p=.03$). The dissimilar-label condition is not included.

There was also an interaction of same onset and same vowel trial types and segment condition, with the model removing this interaction being significantly different than the larger model ($\chi^2=14.53$, $p=.002$). To discover what drove this interaction, we broke it down into a simple interaction analysis by exploring the interaction of trial type for onset conditions and

vowel conditions, respectively, collapsing across merge type. For onset segments, close trials were significantly less accurate than same onset competitor trials ($\beta=-.59$, $SE=.10$, $Z=-5.61$, $p<.001$) and same onset competitor trials were significantly less accurate than same vowel competitor trials ($\beta=-.75$, $SE=.13$, $Z=-5.66$, $p<.001$). For vowel segments, only same onset competitor types were significantly less accurate than same vowel competitor trial types ($\beta=-.22$, $SE=.10$, $Z=-2.15$, $p=.03$).

A few effects were conspicuously absent. First, there were no effects nor interactions with merge type, suggesting that learning phonemically-variable Spanish-based words is no harder than learning words with phonemes that do not vary. Second, and in contrast to the previous experiment, language background also did not have any effects or interactions. That is, Spanish-English bilinguals were neither better nor worse than monolingual English speakers in learning phonotactically-Spanish-like words.

Effects of individual differences on bilingual performance. We also correlated accuracy with the difference between Spanish and English BDS scores calculated earlier. Using Pearson's product-moment correlation, the BDS difference score did not predict a significant correlation between the accuracy data and the difference scores of Spanish minus English on the BDS ($t(38)=.12$, $p=.91$) for the Spanish speaking participants. The relationship was not directional ($r=.02$). We also ran the same type of correlation on scores for scores solely on the Spanish BDS, finding similar results ($t(38)=-.81$, $p=.42$, $r=-.13$). This relationship is shown in Figure 6. The same correlation was made removing the participants in the dissimilar condition, and again, the BDS score did not predict a significant correlation between the accuracy data and the Spanish scores on the BDS ($t(30)=-1.40$, $p=.17$). The relationship was not strongly directional ($r=-.25$).

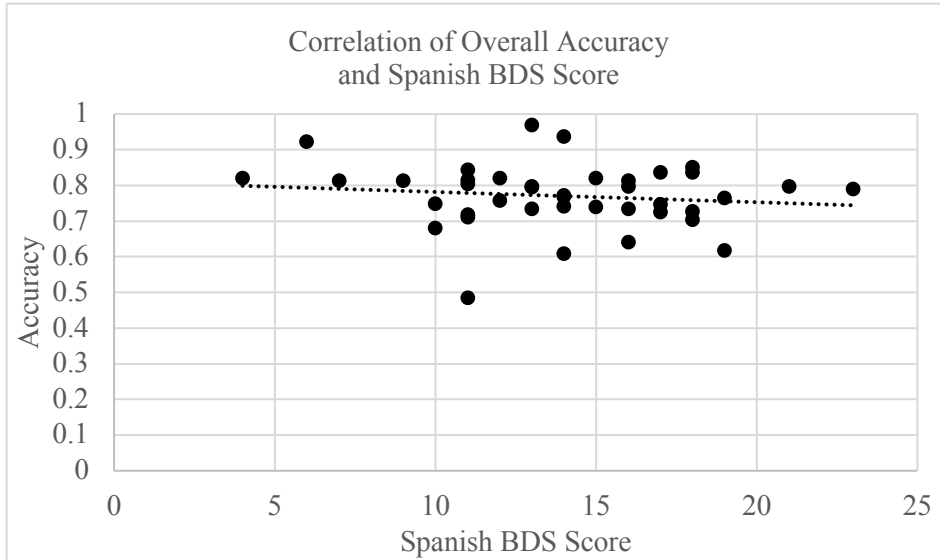


Figure 6. Pearson’s product-moment correlation of overall test accuracy and participants’ score on the Spanish Bilingual Dominance Scale (BDS). There was no significant correlation between Spanish BDS score and accuracy ($p=.42$) nor was it strongly directional ($r=-.13$).

We were also interested in behavior over the learning trials to determine if participants were differing in their trajectories of phoneme merging but eventually arriving at the same accuracy in test. To test this, we built a model that considered merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and whether their training accuracy differed across the first half of the training session in comparison to the second half. We only found a main effect of training half ($\beta=.47$, $SE=.04$, $Z=-11.05$, $p<.001$) with the second half performing better than the first half regardless of merge condition or segment type. Overall, the model that removed training half was significantly different from the larger model ($\chi^2=70.51$, $p<.001$).

Participants were also given a 256-trial two-alternative forced choice discrimination task after the experiment proper to determine that they were able to discriminate the sounds being merged in the experiment. Participants were highly accurate on match trials, with a 14% rate of

false alarms (responding “different” to a “same” trial). On different vowel trials, participants had a 73% hit rate. On different consonant trials, they had a 33% hit rate. On trials with no matching phonemes, they had a 98% hit rate. To explore differences in responses based on trial types, participant condition and trial type were compared using a generalized mixed-effects model with the package lme4 (Bates et al. 2014, R Core Team 2014). Trial types were forward difference coded in terms of false alarm rates; trials with no matching phonemes were compared to trials with same words, trials with the same words were compared to trials with a different vowel, and trials with a different vowel were compared to trials with a different consonant. Participants performed significantly better on totally different trial types than on trial types with matching phonemes ($\beta=2.86$, $SE=.54$, $Z=5.27$, $p<.001$) as well as significantly better on trials with only different vowels than those with different consonants ($\beta=1.72$, $SE=.23$, $Z=7.54$, $p<.001$) (Figure 7). There was no interaction with condition, suggesting that merging of consonants vs. vowels during learning did not change listeners’ underlying perceptual sensitivity.

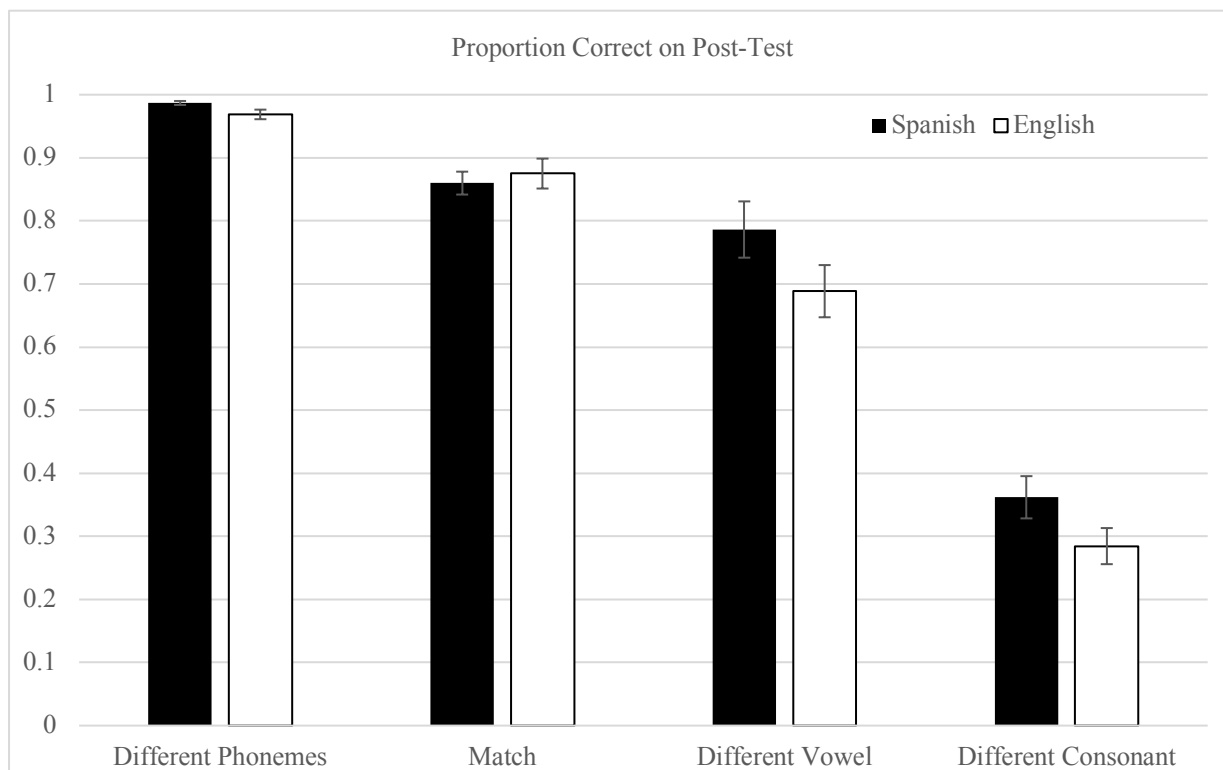


Figure 7. Proportion of correct responses to post-test trials split by language background. When collapsed, accuracy was significantly better for different phoneme trials than different consonant trials ($p < .001$) as well as for different consonant trials versus different vowel trials ($p = .002$).

However, as a whole, participants were still highly accurate (overall error rate: 27%) at distinguishing the sounds being played to them accurately, even if they performed this distinction in the exact opposite manner as requested, providing evidence that an inability to perceive differences between vowels or consonants in the experiment is not what is driving the effects seen in the experiment proper.

To make sure that no differences arose in regard to the different variable phoneme pairs in relation to the native language experience of the participant, we explored a model that took into account language background and consonant merge type. There was no effect of any kind for the different phoneme pairs, even across stops and fricatives, suggesting no difficulty with any variable phoneme in particular, nor any language background receiving any differing effects

on those variable phonemes. This is in spite of the fact that the participants with Spanish-speaking backgrounds tested in this study would have treated some of these sounds as different from the English monolingual participants in day-to-day speech.

Discussion

One important motivation for this study was to disambiguate results from Experiment 1 regarding language background on phoneme merging performance. If native language-specific sub-phonemic information contributes to participants' better performance as seen in Experiment 1, then we would have expected to see the Spanish speakers perform more accurately on this vocabulary.

However, the small advantage of native language background experienced in the first experiment was not evident in this experiment. Nonetheless, this work contributes to the understanding of word learning with respect to bilingualism in other manners. Like previous research (Muench & Creel, 2013), the current experiment found evidence which contradicts numerous other studies on bilingual word learning which contend that bilingualism leads to measurable advantages in word learning. In fact, the background of the participants in the current study did not have a significant effect on any measured aspect of learning the vocabulary. This inconsistency relative to previous studies could be explained in two possible ways. First, language background—and indeed, bilingualism—does not provide participants with measurable benefits when learning small vocabularies in short periods of time. Second, certain aspects of this specific study did not activate bilingual advantages in the way that other tasks might.

Regarding the first explanation, the data from this experiment do not allow us to definitively prove the null hypothesis that there is no bilingual advantage to word learning. However, it is evidence that this advantage is not so wide-reaching that it occurs in any experimental situation. In fact, the combined evidence from this experiment and Experiment 1 demonstrate that being bilingual does not always provide an advantage for word learning tasks, and could even be an inhibiting factor. Further study is required to pinpoint the specific situations that evoke this advantage, as measured in other studies, but this study provides evidence that it is not universal.

To this end, the second explanation gives some insight into what might trigger learning advantages with second languages. In the current experiment, there were no phonemes used from Spanish that do not also exist largely unchanged in English. However, in Experiment 1, as in Muench and Creel's work (2013), there did exist English phonemes (such as /f/) that are not directly comparable to anything in Spanish. This would explain the bilingual advantages found in studies like Kaushanskaya and Marian's (2009) work, where words mismatched English phonology, as well as the difficulties faced by the Spanish-English bilinguals in the experiment 1. This argument would also be supported by work done by Sebastián-Gallés and colleagues on Spanish and Catalan speakers (2009), where Catalan speakers—who have more speech sound categories than Spanish speakers—are capable of representing Spanish sounds as those are a subset of their own sounds. It would also provide an explanation as to why there was no difference according to language background in the current study: all phonemes presented for learning existed in the native languages for both language background groups, so there was no benefit or detriment provided to either group in terms of phonemic familiarity. By contrast, some of the phonemes presented in Experiment 1 were not present in the native language (Spanish) of

many of the bilinguals. It may be that phonemic cues overwhelm the word-learning apparatus, and natively-like subphonemic cues are insufficient to provide a benefit, despite giving native learners meaningful cues to the language being spoken (Grosjean, 1988).

Like Experiment 1, the current study demonstrated again that participants are able to learn merged vocabularies at nearly the same accuracy as control vocabularies, which further speaks to the flexibility of phoneme processing in adults learning novel words. It also demonstrated that providing competition in word recognition can significantly affect accuracy, especially when competitors share multiple phonemes. When targets match competitors in a single phoneme, the merged segment type affects the participant more. For example, a participant in an onset merging condition is arguably ignoring their native distinction between two consonants in order to correctly identify vocabulary they have just learned that ignores those categories. However, when they are faced with a competitor that shares the flexible onset consonant they have been recently taught to allow a bigger category for, this makes them more likely to choose that competitor than if it shared only a vowel with the subject, implying that the participant is still paying attention to the onset consonant, in this case, despite recently learning new categories for it. On the other hand, when a participant is trained on a vocabulary where vowels are merged, the target sharing a vowel with a competitor does not cause the same loss of accuracy. Despite, as in the onset case, the native language barrier between two phonemes being ostensibly lowered, only in cases where the target and competitor share the onset consonant (or the onset consonant as well as the vowel) is there a significant loss of accuracy in either the vowel or the consonant merge condition.

Although these results demonstrate flexibility in adult learners beyond what has been previously considered in the literature, specific types of competition seem to support the primacy

in attention for either the onset position or the consonant over the vowel. This current study cannot differentiate what might be causing the difficulty to learners, but it does further demonstrate the ability of adults to learn words despite phonemic variability even under somewhat difficult test conditions, and adds to evidence that bilingual advantages evident in other domains may be somewhat restricted when it comes to word-learning.

Chapter 3: Differential Phonemic Attributes Effect

When adults perceive spoken utterances, multiple related word candidates are often activated and then vie for eventual comprehension. These lexical relationships can be made in myriad ways: by overlapping with other words (McQueen, Norris, & Cutler, 1994; Tabossi, Burani, & Scott, 1995), overlapping in early material in particular (so-called “cohort” words; Marslen-Wilson & Zwitserlood, 1989), and simply having similar sounding phonemes (Slowiaczek & Hamburger, 1992) or identical phonemes (Radeau, Morais, & Segui, 1995; Slowiaczek, McQueen, Soltano, & Lynch, 2000). These findings of “close but not quite” flexibility in lexical activation raise questions of just how flexible the adult language learning apparatus is when it comes to new vocabulary, and in what conditions modulate that flexibility.

Thus far we have shown adult flexibility in learning not only with variable vowels but also with variable consonants. Our findings of both vowel and consonant flexibility suggest that previous work showing greater “mutability” of vowels in words (Cutler, Sebastian-Galles, Soler-Vilageliu, and Van Ooijen, 2000) may not apply to word learning. What is unclear is whether this flexibility benefits from a privileged syllable (or word) position within tested stimuli, or if this flexibility is broadly inherent to the adult language apparatus. As noted earlier, Creel, Aslin, and Tanenhaus (2006) found evidence that words with shared onset consonants were more confusable than words with shared codas in newly learned vocabulary (see also Creel & Dahan, 2010). A later study with multiple “accents” also found that there was a difference between difficulty of learning coda consonants and nucleus vowels, with consonants causing more difficulty when switched (Muench & Creel, 2013). Muench and Creel, though, did not manipulate the consonant phoneme’s position within the word, nor did they control the degree of phonetic similarity of consonants. This was not the case, however, in Delle Luche et al.’s work

with French and English speakers (2014). Those authors found flexibility for phonemes in a lexical priming task, but they did not find a positional bias. Thus, these two studies conflict; what is driving the consonant flexibility that we have observed in word learning? Does syllable position (onset vs. coda) play a role, or are consonants similarly flexible regardless of syllable position? To answer these questions, we conducted an experiment like the first two, but contrasting the syllable position of the variable consonants. This experiment should definitively discern whether syllable position affects the learning flexibility that has been measured in previous studies. Knowing this will clarify how language learners apply flexibility to newly learned words and improve our understanding of how words are learned in general.

Method

Participants

We tested 80 normal-hearing adults recruited from the UC San Diego SONA subject pool. All participants were native English speakers (avg. age: 22.50 years, sd: 7.36) who grew up hearing no languages besides English and who reported fluency in no other language but English. All participants had less than 3 years of high school language instruction in any other language besides English, and no participant reported exposure to any language besides English before age 14. In addition to the 80 subjects in the data, 24 additional participants were tested, but one was dropped from the final dataset because of an eye-tracking malfunction, seven due to technical difficulties with the post-test, and 16 subjects were replaced in the dissimilar-label condition because the “far” trial types were not comparable in number of phonemes shared to the other conditions.

Stimuli

Visual stimuli were the same as Experiments 1 and 2. Auditory stimuli were a new set of 32 novel consonant-vowel-consonant (CVC) words. The set of consonants used were identical across syllable position, and each consonant was used equally often in each position. While cues to phoneme identity and voicing differ across syllable positions, we deemed it prudent to equate the sound categories at least nominally. Note that to create the desired stimuli we had to resort to using the phoneme /ʒ/ (as in *beige* or *leisure*), which does not typically occur in English syllable onsets. The vocabulary (Table 3) was designed so that 16-pair sets could be made for the merge conditions. These sets were of word pairs which differed only by onset consonant voicing (“*gehb*” and “*kehb*” were one pair), or that differed by coda consonant voicing (“*boog*” and “*book*” were one pair). Words were recorded in a soundproof recording room by one male and one female native English speaker in the Center for Research in Language at UCSD. During the test, only recordings from the male speaker were used. During the post-test, recordings from both speakers were used so that listeners had to respond based on speech sound mismatch rather than acoustic mismatch.

Table 3. A complete list of the nonce words used in the experiment

fehɡ (fɛɡ)	fehʁ (fɛʁ)	buhʃ (bʌʃ)	buhz (bʌz)
vehɡ (vɛɡ)	vehʁ (vɛʁ)	puhʃ (pʌʃ)	puhz (pʌz)
feesh (fiʃ)	feez (fiʒ)	boog (bug)	book (buk)
veesh (viʃ)	veez (viʒ)	poog (pug)	pook (puk)
shuhb (ʃʌb)	shuhp (ʃʌp)	geev (giv)	geef (gif)
zhuhb (ʒʌb)	zhuhp (ʒʌp)	keev (kiv)	keef (kif)
shoov (ʃuv)	shoof (ʃuf)	gehb (geb)	gehp (gɛp)
zhoov (ʒuv)	zhoof (ʒuf)	kehb (keb)	kehp (kɛp)

Design

There were five between-subjects conditions. In the *coda merge* condition, labels for specific objects varied in their coda phonemes (e.g. the same object was labeled both “boog” and “book” /buk/). In the *onset merge* condition, the onset consonant was variable (e.g. “gehb” and “kehb”). In the *dissimilar* (difficult) condition, all phonemes except the vowel varied between the two labels for a single object (e.g. “gehb” and “fehkh”). There were two additional control conditions, where each picture received only one label. The coda control conditions were matched to the coda-merge condition by using either all voiced or all unvoiced consonants present in the vocabulary. The onset consonant control was matched similarly to the onset consonant-merge version in that the control lists used only voiced or unvoiced onset consonants to create the 16-word set. As in Experiments 1 and 2, these control conditions assessed whether reducing the number of functionally different sounds drove effects in the merge conditions, rather than the requirement to learn multiple labels.

Procedure

The design and procedure were the same as Experiment 1, except the learned vocabulary was different. As in Experiments 1 and 2, in close-competitor trials, the target and the competitor shared two of the three phonemes in the word; in the onset consonant-merge and onset consonant-control conditions, target and competitor shared onsets and vowels and differed in codas, e.g., target: /gɛb/-/kɛb/ competitor: /gɛp/-/kɛp/. For the coda-merge and coda-control conditions, target and competitor shared codas and vowels and differed in onsets, e.g., target: /fɛg/-/fɛk/ competitor: /vɛg/-/vɛk/. There was also a same onset and same coda test similarity condition, where target and competitor shared only an onset or coda, respectively (target: /fuv/-

/ʒuv/ competitor: /ʃ^b/-/ʒ^b/; target: /fiʃ/-/fiʒ/ competitor: /b^ʃ/-/b^ʒ/). Finally, in the far competitor test trials, either no sounds were shared or only the coda was shared (target: /b^ʃ/-/b^ʒ/ competitor: /pug/-/puk/ or target: /b^ʃ/-/p^ʒ/ competitor: /fiʃ/-/viʒ/).

Following the experiment proper, participants were given a 256-trial post-test where they were asked to make similarity judgments on two-word pairs. This was to test whether learning merged words led participants to lose perceptual distinctions between similar sounds made by the speaker in the test portion of the experiment. Recall that two native English speakers, a male and a female, recorded the original 32-word vocabulary. Each trial consisted of two spoken words, one from each speaker. Each word in the 32-word vocabulary was played in a word pair an equal number of times, counterbalanced for presentation in the first and second position as well as speaker gender. Participants could respond by typing the letter “g” on the keyboard if the words they heard were the same, and the “0” on the keyboard if they were different. For each word, there were two trials where the exact same word was said by both speakers (once in each of the two possible speaker orders), and six trials where the words were different, in both possible speaker orders. In the different trials, the words could share all sounds but the coda, all sounds but the onset, or no sounds at all.

Finally, all subjects were then administered the MINT test (Gollan et al., 2012) in English to verify baseline fluency in English, and lack of fluency in Spanish. If any subjects mentioned classes in Spanish, the MINT was also delivered in Spanish after the English version. Subjects scored a mean of 63.95 on the English MINT, with a standard deviation of 2.91. Subjects mentioning Spanish class exposure (45 participants) scored a mean of 4.96 on the Spanish MINT, with a standard deviation of 4.41.

Results

The first model assessed whether there was a general effect of label similarity. That is, do we find that learning dual labels is easier when the dual labels are similar to each other, specifically the onset and the coda, than when they are not similar to each other? To test this, as in Experiment 1, we modeled test accuracy in terms of the fixed effects of condition type (onset control, onset merge, vowel control, vowel merge, and hard). Individual subject intercepts and word intercepts and slopes (both fixed effects and their interaction) were included as random effects. The dissimilar-labels learning condition was set as the reference level for this analysis. Each condition type was significantly more accurate than the dissimilar condition (Onset Merge ($\beta=.58$, $SE=.26$, $Z=2.26$, $p=.02$), Coda Merge ($\beta=.74$, $SE=.26$, $Z=2.86$, $p=.004$), Onset Control ($\beta=1.04$, $SE=.26$, $Z=3.99$, $p<.001$), Coda Control ($\beta=.86$, $SE=.26$, $Z=3.31$, $p<.001$)) (Figure 8).

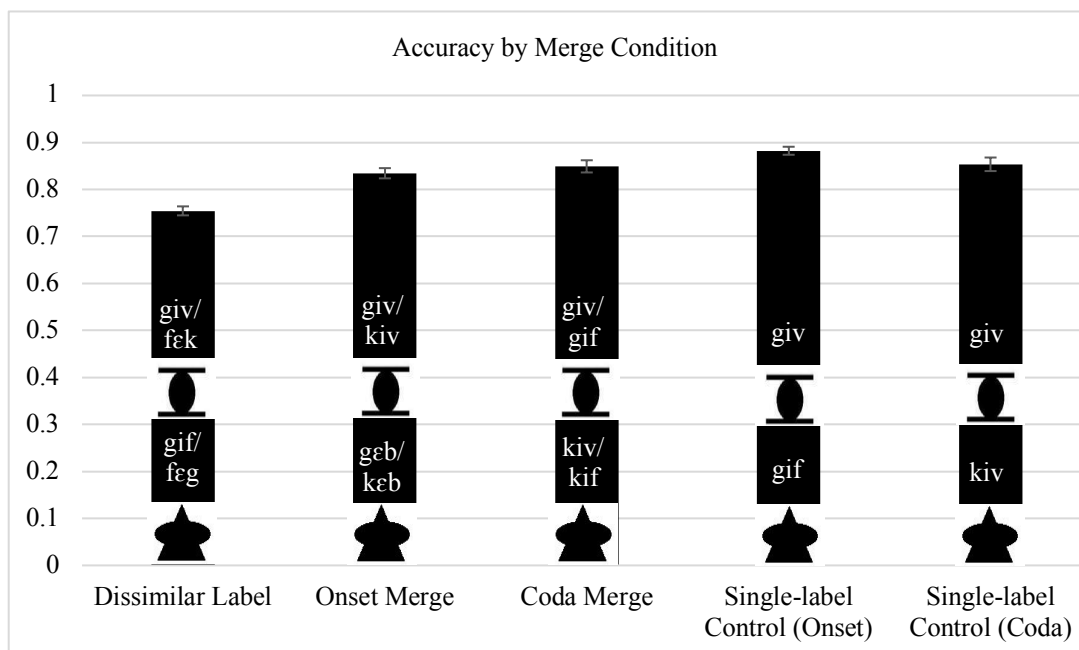


Figure 8. Accuracy in each merge condition. Dissimilar label condition was different from every other condition: Onset Merge ($p=.02$), Coda Merge ($p=.004$), Onset Control ($p<.001$), and Coda Control ($p<.001$).

Removing the Hard condition, we explored a model which considered merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating onset contrasts or coda contrasts), and trial type (either sharing no phonemes, sharing a coda and an onset, sharing a coda and a vowel, or sharing onset and vowel) on participant accuracy. Trial type was forward difference coded, so the contrasts compare trials with successively less-phonologically-similar competitors: close trials to same onset consonant trials, same onset consonant trials to same coda trials, and same coda trials to far trials (no shared phonemes). For this model, there was a main effect of trial type, where the model removing trial type was significantly different from the larger model ($\chi^2=507.77$, $p<.001$). Specifically, accuracy was worse on close competitor trials than same onset ($\beta=-.65$, $SE=.11$, $Z=-5.78$, $p<.001$), and same onset competitor trials were significantly less accurate than same coda competitor trials ($\beta=-1.29$, $SE=.16$, $Z=-7.96$, $p<.001$). See Figure 9.

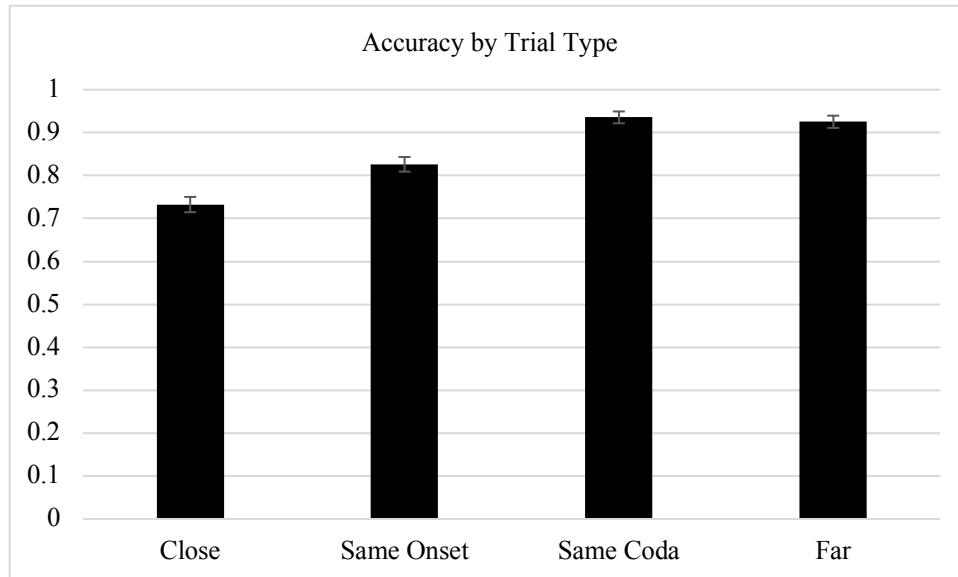


Figure 9. Close trials were significantly different from same onset trials ($p < .001$), which were significantly different from same vowel trials ($p < .001$). The dissimilar-label condition is not included here.

There was also an interaction of same onset trial type and segment condition ($\beta = -.39$, $SE = .10$, $Z = -3.85$, $p < .001$). Exploring this interaction further, we built models of trial type in onset conditions and coda conditions. In the onset condition, same onset competitor trials ($\beta = -1.18$, $SE = .19$, $Z = -6.31$, $p < .001$) had accuracy that was higher than on close competitor trials, and same coda trial types ($\beta = -2.13$, $SE = .47$, $Z = -4.50$, $p < .001$) had accuracy that was higher than same onset competitor trials. In the coda condition, there was only an effect of same coda competitor trials vs. same onset trials ($\beta = -2.14$, $SE = .36$, $Z = -5.90$, $p < .001$), with accuracy on same coda trials being higher.

There was also an interaction of trial types and merge condition. As above, we explored this interaction by building a simple analysis of trial types on merged vocabularies and control vocabularies. In merged vocabularies, differences appeared for certain trial types. Specifically, close competitor trial types were less accurate than same onset competitor trials ($\beta = -.85$, $SE = .13$,

$Z=-6.41$, $p<.001$), which were subsequently less accurate than same coda trials ($\beta=-1.54$, $SE=.22$, $Z=-7.16$, $p<.001$), which were less accurate than far competitor trials ($\beta=.53$, $SE=.24$, $Z=2.22$, $p=.03$).

On control vocabularies, a slightly different pattern appeared. First, same onset competitor trials ($\beta=-.53$, $SE=.24$, $Z=-2.20$, $p=.03$) were more accurate than close competitor trials, but less accurate than same coda competitor trials ($\beta=-1.74$, $SE=.37$, $Z=-4.75$, $p<.001$).

Importantly, despite these interactions, there was no main effect of Merge Type, nor a Merge Type by Segment Type interaction. This means that, overall, merging onset consonants or coda consonants appeared to have no detrimental effects on accuracy.

We were also interested in behavior over the learning trials to determine if participants were differing in their trajectories of phoneme merging but eventually arriving at the same accuracy in test. To test this, we built a model that considered merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and whether their training accuracy differed across the first half of the training session in comparison to the second half. By default, this removed the dissimilar condition. We found a main effect of training section ($\beta=.69$, $SE=.07$, $Z=9.43$, $p<.001$). There was also an interaction of merge condition and segment condition, but it was qualified by a three-way interaction of merge condition, segment type, and section half ($\beta=-.20$, $SE=.07$, $Z=-2.72$, $p=.007$).

To discover what drove this interaction, we broke it down into two simple interaction analyses at each level of section half. First, we explored the interaction of merge type and segment type for the first half, and second that same interaction for the second half of training trials. For first half training trials there was no main effect. For second half training trials there

was a main effect of merge condition ($\beta=.45$, $SE=.23$, $Z=1.97$, $p=.05$) as well as an interaction between merge type and segment type ($\beta=-.52$, $SE=.22$, $Z=-2.34$, $p=.02$). We further broke this interaction down into a single interaction of segment type during the second half of training trials for merged phonemes and control phonemes. There was only an effect for the control phonemes ($\beta=-.91$, $SE=.37$, $Z=-2.46$, $p=.01$), implying by the second half of the training trials, participants in control vocabularies are significantly less accurate on the coda consonants than the onset consonants.

Eye tracking analyses. Participants were also eye-tracked during the learning and testing portions of this experiment. To explore the participants' visual search patterns and assess whether participant looks corroborated their accuracy data, eye-tracking data from the testing portion of the experiment were analyzed with mixed-effects models. Regression was conducted following Barr (2008) and was done separately with participants and items as random effects to accrue multiple trials per data point.

In the first eye-tracking model we assessed the effect of trial type and test condition on the empirical-logit-transformation of looks to the target (Barr, 2008). Looks were analyzed in a 1-s (1000-ms) time window ranging from 200 to 1200ms to encompass the full length of the single word. ANOVAs were run in both a by-subject and by item analysis. For the subjects and word analysis, there were main effects of trial type (Subjects: ($F(3, 222)=68.50$, $p<.001$); Words ($F(3, 93)=37.13$, $p<.001$) and condition (Subjects: ($F(4, 74)=4.53$, $p=.002$); Words ($F(4, 124)=12.37$, $p<.001$)). There was also an interaction in the subject analysis of trial type and condition ($F(12, 222)=2.07$, $p=.02$), but it did not occur in the words analysis and was not explored.

The main effect of trial type is explained in the subject analysis by the close trials receiving significantly fewer looks than the same onset trials ($t(78)=-3.96, p=.001$), the same onset trials receiving fewer looks than the same coda trials ($t(78)=-9.02, p<.001$), but the same coda trials not being different from the far trials ($t(78)=1.59, p=.12$). This same pattern occurred in the word analysis, with close trials receiving significantly fewer looks than the same onset trials ($t(31)=-2.87, p=.005$), the same onset trials receiving fewer looks than the same coda trials ($t(31)=-8.50, p<.001$), but the same coda trials not being different from the far trials ($t(31)=0.93, p=.35$). The main effect of condition demonstrates in the subject and word analysis that all conditions except the merging of onsets received significantly more looks than the hard condition (Subjects: onset control $F(1, 30)=16.53, p<.001$, onset merge $F(1, 30)=1.47, p=.23$, coda control $F(1, 30)=12.26, p=.001$, coda merge $F(1, 30)=8.61, p=.006$; Words: onset control $F(1, 223)=23.09, p<.001$, onset merge $F(1, 223)=2.04, p=.16$, coda control $F(1, 223)=27.13, p<.001$, coda merge $F(1, 223)=13.06, p<.001$).

A second ANOVA on eye-tracking data was performed after removing the hard condition data. In the second eye-tracking model we assessed the effect of trial type, merge condition, and segment type on the empirical-logit-transformation of looks to the target (Barr, 2008). Looks were again analyzed in a 1-s (1000-ms) time window ranging from 200 to 1200ms to encompass the full length of the single word. ANOVAs were run in both a by-subject and by item analysis. In these models, there were main effects of merge type (Subject: $F(1, 59)=4.54, p=.04$; Word: $F(1, 31)=12.24, p=.001$) and trial types (Subject: $F(3, 177)=57.37, p<.001$; Word: $F(3, 93)=31.80, p=.001$). In the word analysis, there was a main effect of segment type ($F(1, 31)=5.30, p=.03$), but it did not occur in the subject analysis so it was not explored. There was also a three-

way interaction of trial type, merge type, and segment type in the subject analysis ($F(3, 177)=3.12, p=.03$), but it was not in the word analysis so it was not explored.

The main effect of trial type is explained in the subject analysis by the close trials receiving significantly fewer looks than the same onset trials ($F(1, 62)=9.02, p=.004$), the same onset trials receiving fewer looks than the same coda trials ($F(1, 62)=78.03, p<.001$), but the same coda trials not being different from the far trials ($F(1, 62)=0.86, p=.35$). This same pattern occurred in the word analysis, with close trials receiving significantly fewer looks than the same onset trials ($F(1, 31)=4.58, p=.04$), the same onset trials receiving fewer looks than the same coda trials ($F(1, 31)=56.94, p<.001$), but the same coda trials not being different from the far trials ($F(1, 31)=0.16, p=.69$). The main effect of merge condition demonstrates in the subject and word analysis that merged conditions received significantly fewer looks than the control conditions (Subjects: $F(1, 61)=4.55, p=.04$; Words: $F(1, 479)=8.81, p=.003$).

Participants were also given a 256-trial two-alternative forced choice discrimination task after the experiment proper to determine that they were able to discriminate the sounds being merged in the experiment. Participants were highly accurate on match trials, with a 6% rate of false alarms (responding “different” to a “same” trial). On different coda trials, participants had a 76% hit rate. On different onset trials, they had a 74% hit rate. On trials with no matching phonemes, they had a 98% hit rate. To explore differences in responses based on trial types, participant condition and trial type were compared using a generalized mixed-effects model with the package lme4 (Bates et al. 2014, R Core Team 2014). Trial types were forward difference coded in terms of false alarm rates; trials with no matching phonemes were compared to trials with same words, trials with the same words were compared to trials with a different coda, and trials with a different coda were compared to trials with a different onset. Participants performed

significantly better on totally different trial types than on trial types with all matching phonemes ($\beta=1.45$, $SE=.31$, $Z=4.66$, $p<.001$) as well as significantly better on trials with all matching phonemes than those with different codas ($\beta=1.61$, $SE=.18$, $Z=8.94$, $p<.001$) (Figure 10). There was also interaction of different coda ($\beta=.12$, $SE=.05$, $Z=2.27$, $p=.02$) and different onset ($\beta=.09$, $SE=.04$, $Z=2.32$, $p=.02$) with condition.

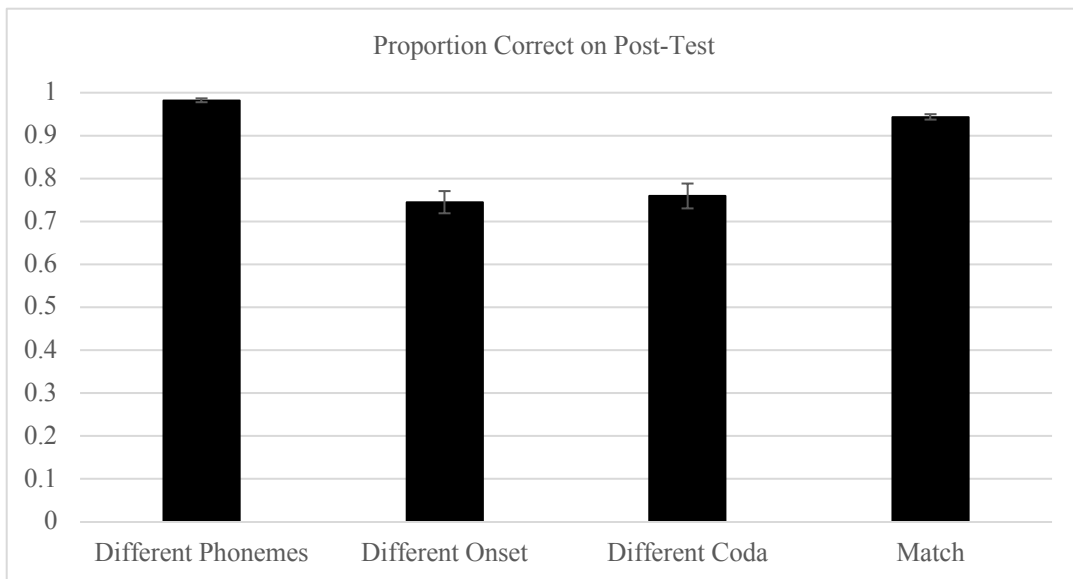


Figure 10. Proportion of correct responses to post-test trials was significantly better for different phoneme trials than different consonant trials ($p>.001$) as well as for different vowel trials ($p>.001$).

To explore this interaction, we built a simple analysis of the trial types on each condition. For those trained in the Hard condition, participants performed significantly worse on trials with totally different phonemes than those with entirely matching phonemes ($\beta=.74$, $SE=.25$, $Z=2.97$, $p=.003$) and significantly worse on trials where competitor and target had different codas than those with all matching phonemes ($\beta=1.95$, $SE=.17$, $Z=11.42$, $p<.001$). For the coda merge condition, participants were significantly better at trials with matching phonemes than trials with

all different phonemes ($\beta=6.93$, $SE=2.36$, $Z=2.94$, $p=.003$). Participants in the onset merge condition performed significantly worse on trials with matching words than those with totally different phonemes ($\beta=1.04$, $SE=.44$, $Z=2.37$, $p=.02$), and significantly better on trials with different codas than trials with all matching phonemes ($\beta=2.14$, $SE=.59$, $Z=3.62$, $p<.001$). However, as a whole, participants were still highly accurate (overall error rate: 14%) at distinguishing the sounds being played to them accurately, providing evidence that an inability to perceive differences between onset or coda consonants in the experiment is not what is driving the effects seen in the experiment proper.

Discussion

In this study, we asked whether learners had greater difficulty learning words with variable onset consonants vs. variable coda consonants. At a coarse grain, we expected the results of this study to mimic those of the first study in this document. That is to say, for English speakers learning English-based vocabularies, vocabularies merging dissimilar words would be the most difficult, single-word (control) vocabularies would be the least difficult, and vocabularies that merged vowels and consonants would fall between those endpoints. We also expected the eye-tracking results to mimic the accuracy findings. As in the first two experiments, the dissimilar (hard) label merging was most difficult, and again, vocabularies that merged onsets or codas were no more difficult than matched control vocabularies that taught only single labels. Based on work mentioned previously as well as the first two experiments in this document, it is clear adults can quickly learn similar phonemically-variable vocabularies with little detriment to accuracy. Nonetheless, eye tracking data showed that recognition was somewhat slower for phonemically-variable vocabularies than control vocabularies.

Where this experiment adds to our understanding of phonemic flexibility is in the interactions of the segment conditions with the trial types. Like experiments 1 and 2, trials where competitors shared multiple phonemes were significantly more difficult than were trials sharing only a single phoneme. Also as in experiments 1 and 2, the participants run into the most trouble when the competitor shares an onset phoneme with the target. The later the phoneme appears in a word, the less stringent the learner seems to be with the phonemic boundaries. The vast majority of studies that argue consonants are more categorical in nature rely on onset consonant words, and subsequently there has been a general agreement that consonants are more categorical and less amenable to flexibility. These data point to the interpretation that at least when challenged with mildly difficult competitor words, this widely-agreed upon outcome is dependent on the position of the phoneme tested.

Most tellingly, participants struggled on accurately choosing the target in those trials where the competitor was sharing an onset compared with those trials where the competitor is sharing a coda. If phoneme merging is performed on a phoneme-type-wide basis, that would imply that accuracy in the face of competitors sharing onsets and those sharing codas should be the same. This onset consonant bias has been discussed before (Creel, Aslin, & Tanenhaus, 2006; Creel & Dahan, 2010), and would predict, if the important aspect was the type of phoneme (that is, consonant vs. vowel), that merging consonants elsewhere in the word would provoke similar problems. However, what we actually see is an effect of position, where consonants that are merged later in the word cause less confusion when given the same competitors. This is not to say that no confusion is caused by merging the phonemes and providing competing objects, but the effects are mitigated by the position of the merge.

Chapter 4: Phonemic Merge in Search Space Effect

We know that it is possible to merge certain phonemes in behavioral settings. What is unclear is how deeply this merging affects low-level perception. Does merging certain phonemes affect how images are searched in a more real-world paradigm? Using eye-tracking, we can gain insight into the pre-decision processes that occur in word learning before behavioral responses occur (Tanenhaus et al., 1995). In this manner, we can observe linguistic processing that slower and more deliberate behaviors, such making a response selection, might tend to mask. Specifically, using visual fixations to target pictures as a dependent measure will allow us to monitor responses to specific phonemes as they are processed by the auditory system, but before they are fully comprehended and acted upon with a programmed hand movement.

If the sound presented in the critical trial has been merged perceptually, predictive eye movements will be recorded toward the correct choice in the trial at the same speed as for participants in the single-word conditions. This experiment will provide greater insight into the depth of processing when learning a new vocabulary as well as demonstrate the speed of flexibility in processing merged phonemes. It will also speak to the extent to which, if listeners do actually perceptually alter speech sound representations, they are merged in particular word contexts versus more generally without respect to word contexts.

Method

Participants

We tested 80 normal-hearing adults recruited from the UC San Diego SONA subject pool. All participants were native English speakers (avg. age: 21.35 years, sd: 9.36) who grew up hearing no other languages besides English and reported fluency in no other language but

English. All participants had less than 3 years of high school language instruction in any other language besides English, and no participant reported exposure to any language besides English before age 14. In addition to the 80 subjects used in the data, 25 additional participants were tested but dropped from the final dataset and replaced: five because of equipment failures moving from the test to the post-test, four due to a coding error in one condition that was discovered during the experiment, and 16 subjects in the Vowel Change condition because the target objects all occurred on the same side of the screen for a particular test trial type, affecting the eye-tracking data.

Stimuli

Visual stimuli were a set of 16 two-dimensional monochromatic nonsense shapes first used by Creel et al. (2006). Auditory stimuli were the same 32 novel consonant-vowel-consonant (CVC) words as used in Experiment 1.

Procedure

The design and procedure were the same as Experiment 1 with the exception that training and test trials were eye tracked. During each training or testing trial, participants' eye movements were tracked using an Eyelink 1000 Desktop eye tracker. Calibration accuracy was checked between trials via a drift-correction event.

Following the experiment proper, participants were given a 256-trial post-test where they were asked to make similarity judgments on two-word pairs. This was to test whether participants had become insensitive to the perceptual contrasts that were variable in the test portion of the experiment. Two native English speakers, a male and a female, recorded the

original 32-word vocabulary. Each trial consisted of two spoken words, one from each speaker. The voice change required listeners to make a phonetic match rather than an exact acoustic match. Each word in the 32-word vocabulary was played in a word pair an equal number of times, counterbalanced for presentation in the first and second position as well as speaker gender. Participants could respond by typing the letter “g” on a keyboard if the words they heard were the same, and the “0” on the number pad if they were different. For each word, there were two trials where the exact same word was said by both speakers (once in each of the two possible speaker orders), and six trials where the words were different, in both possible speaker orders. In the different trials, the words could share all sounds but the vowel, all sounds but the onset, or no sounds at all.

Finally, all subjects were then administered the MINT test (Gollan et al., 2012) in English to verify baseline fluency in English. If any subjects mentioned classes in Spanish, the MINT was also delivered in Spanish after the English version. Out of 68 possible points, subjects scored a mean of 64.13 on the English MINT, with a standard deviation of 2.14. English monolingual subjects mentioning Spanish class exposure (33 participants) scored a mean of 6.12 on the Spanish MINT, with a standard deviation of 4.69. The highest score achieved by any participant was 16.

Results

Like other experiments, the first model assessed whether there was a general effect of label similarity, but this experiment was novel (along with Experiment 3 in the previous chapter) in that we also monitored eye tracking. That is, do we find that learning dual labels is easier when the dual labels are similar to each other than when they are not, and are there effects

evident in a more sensitive measure such as eye tracking? To test this, we modeled test accuracy in terms of the fixed effects of condition type (onset control, onset merge, vowel control, vowel merge, and hard). Individual subject intercepts and word intercepts and slopes (both fixed effects and their interaction) were included as random effects. The dissimilar-labels learning condition was set as the reference level for this analysis. Each condition type was significantly more accurate than the dissimilar condition (Onset Merge ($\beta=1.71$, $SE=.25$, $Z=6.76$, $p<.001$), Vowel Merge ($\beta=2.12$, $SE=.27$, $Z=7.73$, $p<.001$), Onset Control ($\beta=1.71$, $SE=.25$, $Z=6.76$, $p<.001$), Vowel Control ($\beta=2.63$, $SE=.29$, $Z=9.13$, $p<.001$)) (Figure 11).

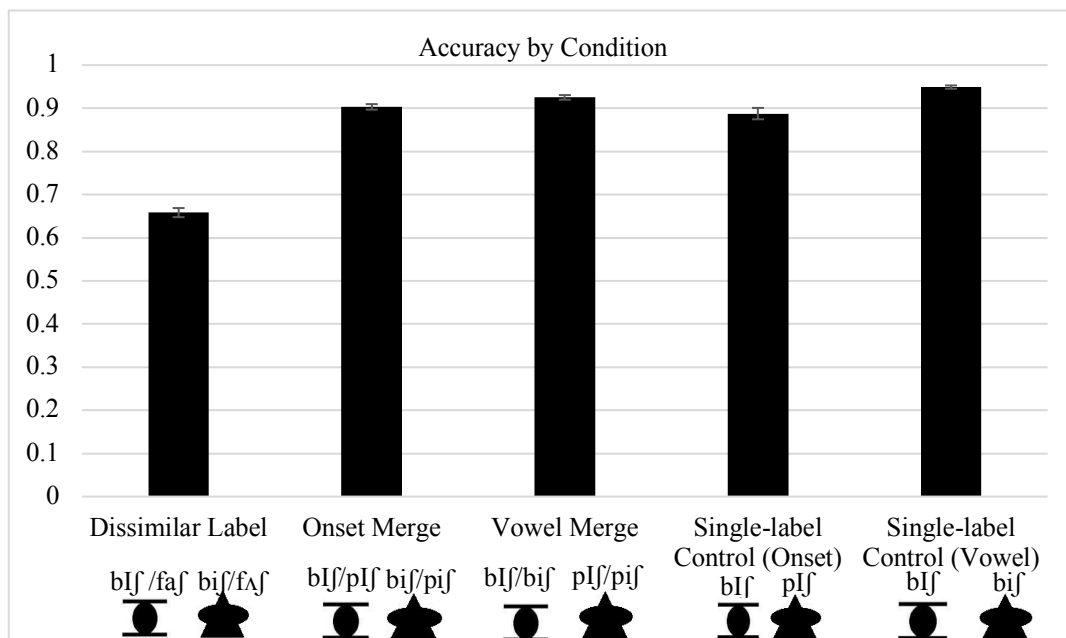


Figure 11. Accuracy in each merge condition. Dissimilar label condition was different from every other condition, $p<.001$.

A second model asked more specific questions regarding the effect of merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts

or vowel contrasts), and trial type (either sharing no phonemes, sharing a coda and an onset, sharing a coda and a vowel, or sharing onset and vowel) on participant accuracy. Again, trial type was forward difference coded, so the contrasts compare trials with successively less-phonologically-similar competitors: close trials to same onset consonant trials, same onset consonant trials to same vowel trials, and same vowel trials to far trials (no shared phonemes). The dissimilar labels condition was dropped for this model. There was a main effect of trial type, with accuracy in close competitor trials being worse than in same onset competitor trials ($\beta=-1.03$, $SE=.19$, $Z=-5.48$, $p<.001$), accuracy in those trials being worse than in same vowel competitor trials ($\beta=-1.31$, $SE=.32$, $Z=-4.02$, $p<.001$), and accuracy in those trials being worse than in far competitor trials ($\beta=-1.28$, $SE=.51$, $Z=-2.49$, $p=.01$) (See Figure 12 for details).

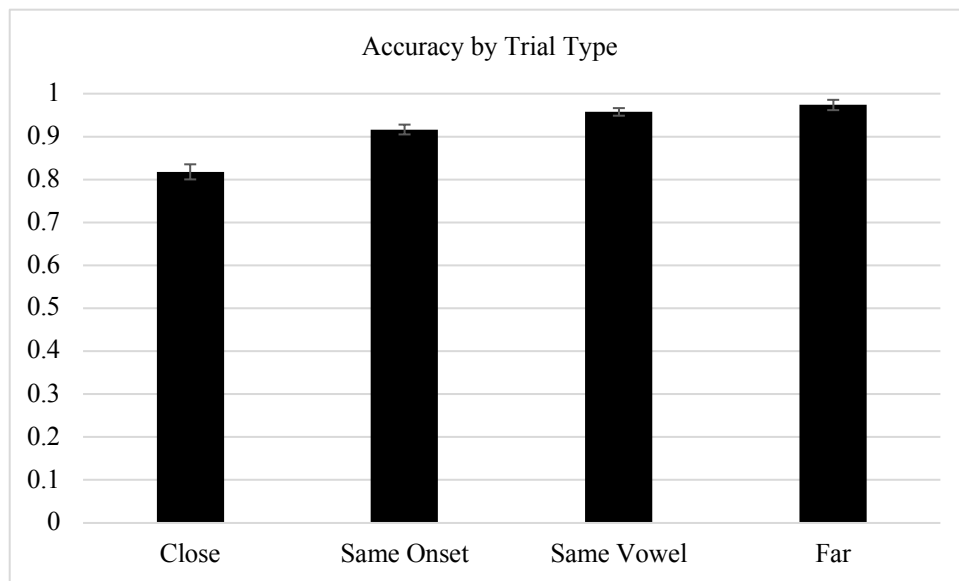


Figure 12. Close trials were significantly different from all same consonant trials ($p<.001$), which were significantly different from same vowel trials ($p<.001$), which were significantly different from far trials ($p=.01$). The dissimilar-label condition is not included here.

The model that removed this interaction was significantly different from the larger model ($\chi^2=328.56$, $p<.001$).

There was also an interaction of trial type and segment condition, with the model that removed this interaction being significantly different than the larger model ($\chi^2=85.87$, $p<.001$). Exploring this interaction further, we built models of trial type in onset conditions and vowel conditions, collapsing over merge type. For the onset condition, there were main effects of trial type, with accuracy being significantly worse in close competitor trials than same onset competitor trials ($\beta=-2.03$, $SE=.28$, $Z=-7.28$, $p<.001$) and those same onset trials being significantly worse than same vowel competitor trials ($\beta=-.87$, $SE=.44$, $Z=-2.02$, $p=.04$).

In the vowel conditions, there was only a main effect of trial type for the difference between same onset competitor trials and same vowel competitor trials ($\beta=-1.58$, $SE=.47$, $Z=-3.39$, $p<.001$), with performance on the latter trial types being significantly better than on the former.

Importantly, there was no effect of merge type, nor did it interact with other factors, suggesting that learning vocabularies with variable sounds does not impair accuracy relative to learning with invariant sounds.

We were also interested in behavior over the learning trials to determine if participants were differing in their trajectories of phoneme merging but eventually arriving at the same accuracy in test. To test this, we built a model that considered merge condition (whether a subject was taught a vocabulary with merged sounds, or the control form with only one sound), segment type (whether the vocabulary was based on eliminating consonant contrasts or vowel contrasts), and whether their training accuracy differed across the first half of the training session in comparison to the second half. We found a main effect of segment type ($\beta=.25$, $SE=.10$,

$Z=2.40, p=.02$) and training section ($\beta=.80, SE=.05, Z=15.31, p<.001$). There was also an interaction of segment type and section half ($\beta=.10, SE=.05, Z=2.05, p=.04$).

To discover what drove this interaction, we broke it down into two simple interaction analyses at each level of section half. For first half training trials there was no main effect of segment type ($\beta=.14, SE=.08, Z=1.82, p=.07$). For second half training trials there was a main effect of segment type ($\beta=.37, SE=.17, Z=2.22, p=.03$) with onset trials (94.8%) having higher accuracy than vowel trials (89.3%).

Participants were also eye-tracked during the learning and testing portions of this experiment. To explore the participants' visual search patterns and assess whether participant looks corroborated their accuracy data, eye-tracking data from the testing portion of the experiment were analyzed with ANOVAs. Data were empirical-logit transformed following Barr (2008), separately with participants and items as random effects to accrue multiple trials per data point. (Measuring looking proportion on a single trial results in a distribution that is neither binomial nor normal; aggregating across trials yields a distribution that is closer to normal.)

In the first eye-tracking model we assessed the effect of condition and test trial type on the empirical-logit-transformation of looks to the target (Barr, 2008). Looks were analyzed in a 1-s (1000-ms) time window ranging from 200 to 1200ms to encompass the full length of the single word. ANOVAs were run in both a by-subject and by item analysis. For the subjects and word analysis, there were main effects of trial type (Subjects: ($F(3, 222)=29.59, p<.001$); Words ($F(3, 93)=34.73, p<.001$) and condition (Subjects: ($F(4, 74)=5.96, p<.001$); Words ($F(4, 496)=36.64, p<.001$), and an interaction between the two (Subjects: ($F(12, 222)=4.53, p<.001$); Words ($F(12, 496)=4.72, p<.001$).

The main effect of condition demonstrates in the subject and word analysis that all conditions received significantly more looks than the hard condition (Subjects: onset control $F(1,29)=15.28, p<.001$, onset merge $F(1, 32)=19.68, p<.001$, vowel control $F(1, 32)=13.20, p<.001$, vowel merge $F(1, 32)=24.25, p<.001$; Words: onset control $F(1, 223)=49.94, p<.001$, onset merge $F(1, 223)=65.86, p<.001$, vowel control $F(1, 223)=96.17, p<.001$, vowel merge $F(1, 223)=86.34, p<.001$).

A second ANOVA on eye-tracking data was performed after removing the hard condition data. In the second eye-tracking model we assessed the effect of trial type, merge condition, and segment type on the empirical-logit-transformation of looks to the target (Barr, 2008). Looks were again analyzed in a 1-s (1000-ms) time window ranging from 200 to 1200ms to encompass the full length of the single word. ANOVAs were run in both a by-subject and by item analysis. In these models, there was a main effect of trial type (Subject: $F(1, 57)=25.17, p<.001$; Word: $F(3, 93)=31.24, p<.001$) as well as a two-way interaction between trial type and segment condition (Subjects: $F(3, 171)=11.22, p<.001$; Words $F(3,372)=11.30, p<.001$). In the word analysis, there was a main effect of segment condition as well ($F(1, 372)=3.99, p=.05$), but it did not occur in the subject analysis so it was not explored. There were no effects of Merge Type or interactions, suggesting that inconsistent sound patterns did not detectably affect visual fixations. The interaction between the trial types and the segment condition was explored by examining the effect of segment condition on each test trial type. The only trial type affected by the segment condition, in both the subject and word analyses, was the close test trial (Subject: $F(1, 59)=18.21, p<.001$; Word: $F(1, 95)=28.39, p<.001$). Subsequent analyses found that in the close trial, there were significantly fewer looks to targets for onset conditions than for vowel conditions, in both subject and word analyses (Subject: $t(59)=-4.27, p<.001$; Word: $t(31)=-5.08,$

$p < .001$). The main effect of trial type is explained in the onset condition by the close trials receiving significantly fewer looks than the same onset trials (subjects: $t(28) = -6.29, p < .001$; items: $t(31) = -6.54, p < .001$) and the same vowel trials receiving significantly fewer looks than the far trials (subjects $t(28) = -2.03, p = .05$; items: $t(31) = -2.13, p = .04$). In the vowel condition of the subject analysis, only the same onset trials received significantly fewer looks than the same vowel trials ($t(31) = -2.61, p = .01$). This same pattern occurred in the vowel condition of the word analysis, with only the same onset trials receiving significantly fewer looks than the same vowel trials ($t(31) = -3.09, p = .003$).

Due to experimenter error, only 6 participants were also given a 256-trial two-alternative forced choice discrimination task after the experiment proper to determine that they were able to discriminate the sounds being merged in the experiment. Five were participants in the vowel merge condition, and one was in the onset consonant control condition. Participants were highly accurate on match trials, with a 4% rate of false alarms (responding “different” to a “same” trial). On different vowel trials, participants had a 56% hit rate. On different consonant trials, they had a 76% hit rate. On trials with no matching phonemes, they had a 94% hit rate. Importantly, participants performed no better on trial types with different consonants than on trial types with different vowels ($t(5) = 1.56, p = .18$). Performance on different vowel trials being worse than chance for multiple participants indicates directions to treat word pairs in the post-test as separate from any learning in the test proper were ignored. However, as a whole, participants were still highly accurate (overall false alarm rate: 20%) at distinguishing the sounds being played to them accurately, supporting the indication that an inability to perceive differences between vowels or consonants in the experiment is not what is driving the effects seen in the experiment proper.

Discussion

We expected this study to mirror the results of prior experiments regarding participant accuracy on word learning because the stimuli and procedure were the same with the addition of the participants being eye-tracked. This means that dissimilar vocabularies would be more difficult than all other merged vocabularies, and that merged vocabularies would be similarly learnable to their single-label control conditions. Our results indicate that this same trend, noted in Experiments 1-3, did in fact occur. What is most interesting about the pattern of results from this specific experiment is its close match to the pattern in Experiment 3.

Again, it was the differing trial types in the testing portion of the experiment that demonstrated a significant change in difficulty for participants. When sharing multiple phonemes, participants were significantly more likely to get confused and make the wrong choice in a two-alternative forced choice task than when competitor objects shared only one or no phonemes. This demonstrates again that learning merged vocabularies is well within the purview of capability for adult learners, further challenging the notion that phonemes exist as discretely learned linguistic building blocks. It also argues that early in the process of learning a new vocabulary, the flexibility of these phonemes is susceptible to confusion when enough phonemes are shared between words.

This difficulty in accurately recognizing a target instead of a close competitor also interacted with the segment type that was being merged. Most interestingly, this experiment followed the exact same pattern of competitor influence on accuracy by segment location. The increased importance of the onset consonant has been demonstrated in multiple experiments within this document and in other studies (Creel, Aslin, & Tanenhaus, 2006; Creel & Dahan, 2010), and in conjunction with other results from this document add more support to a positional

importance to the ease phonemic merging when learning a novel vocabulary. That is to say, when the onset consonant is merged, competitors sharing anything beside the coda will change accuracy, but when the vowel— a phoneme in the second position in the word— is merged, only competitors that directly share that phoneme affect accuracy. Further exploration of vocabularies that start with vowels are necessary to fully support this theory, but the current findings do provide compelling evidence to explore it.

Chapter 5: Developmental Effects on Phonemic Merging

As children develop, one of the most crucial activities they take part in is language acquisition. One of the hallmarks of linguistic maturation is becoming a better communicator and demonstrating adult-like speech. However, the details of this process are not well understood, nor is the timeline for maturational changes well delineated. Language learners are exposed to a wide range of sounds when learning words, and tokens can differ by aspects like accent or one's region of origin (Eckert, 2008; Labov, Ash, & Boberg, 2006), and prosody. Despite marked acoustic differences making recognition of the same word form across different speakers and contexts somewhat difficult for infants (e.g. Houston & Jusczyk, 2000), children are nonetheless consistently able to acquire new words.

How does this occur? One heuristic that monolingual children are thought to apply is that of mutual exclusivity; when one or more objects with known label(s) and an object without a known label are displayed, and a label that doesn't describe the known object is said, the new object must apply to the new label (Markman & Wachtel, 1988, Golinkoff, Mervis, & Hirsh-Pasek, 1994; Halberda, 2003). Due to how sensitive we know children are to specific phonemes within their language (Werker & Tees, 1984), recognizing new words should be a matter of detecting that one is hearing a novel combination of phonemes. If children are constantly on the alert for novel sequences of phonemes, then it should be highly challenging to learn a single word with variable phonemes.

However, there is evidence that children are more flexible in their phoneme perception during word learning than a phonemes-as-symbols account would allow for. For instance, in a study by Creel (2012), 3-5-year-old children who were asked to point to a picture of a "fesh" pointed to a fish picture significantly more often than to a novel picture, implying that they were

parsing this vowel change as simply another version of the word “fish.” This performance implies some flexibility in the use of mutual exclusivity and categorical processing accounts. There is evidence as well that learning phonologically inconsistent tokens of a word does not necessarily mean learning many forms separately. Muench & Creel (2013) found that when variable vowels within a word differed only slightly (e.g. /i/-/I/), adults were just as quick to learn twice as many labels for objects, implying that the process to learn multiple label to object mappings was the same as learning a single label to object mapping. Is this flexibility learned very slowly, or is it available during development?

How flexible are children when it comes to sound to meaning mappings? The current study uses classic behavioral response measures (pointing accuracy) in combination with eye-tracking methods to explore the range of flexibility in phoneme learning in preschool aged children. It probes learning difficulty for merging similar but categorically different vowels and onset consonants. As some earlier literature has demonstrated, adults find learning multiple labels with similar—but not identical—vowels for a single object to be just as easy as learning a single label for a single object, with consonants somewhat more difficult to merge. However, will children show the same flexibility as adults, or a different pattern of flexibility, particularly given that children’s consonant perception does not appear fully adult-like until perhaps after age 12 (Hazan & Barrett, 2000)? We predict that while children may be capable of learning merged-label objects, the immature nature of their phoneme processing means they should not be significantly better at merging vowels or consonants specifically.

Method

Participants

Sixty-four monolingual English-speaking preschool children ages 3-5 years (average: 4.34, SD: .70) participated. Thirty-two participated in the vowel merge condition, and 32 participated in the consonant merge condition. Ten additional participants were tested but dropped from the final dataset because they did not finish the experiment due to fussiness, 3 were dropped for technical issues, and 2 for excess exposure to languages other than English.

Stimuli

Visual stimuli are child-friendly cartoons created in the lab, as well as 4 colorful high-resolution images used for getting the child's attention (a crawling baby, a kitten, a panda, and a deer). Auditory stimuli were the 32 novel consonant-vowel-consonant (CVC) words used in Chapter 1 and Chapter 4.

Procedure

The study consisted of teaching eye-tracked participants to recognize two cartoon characters per block, followed by two-alternative testing on character recognition using pointing (no feedback). Each participant completed 2 blocks of training and testing with two different sets of cartoon characters (one train-test cycle was followed by a second train-test cycle). During each test trial, participants' eye movements were tracked using an Eyelink 1000 Desktop Eyetracker. Calibration accuracy was checked between trials via a drift-correction event.

In one block, children learned a single label for each of two creatures (creature one: "deev", creature two: "zuf"). In another block, they learned a merged word pair for creatures

(either a vowel-merged version, creature one: faʃ/faʃ, creature two: deɪdʒ/dɛdʒ; or a consonant merge condition, creature one: faʃ/vaʃ, creature two: deɪdʒ/teɪdʒ). Block order was counterbalanced across participants. Half of the children were assigned to the consonant-merge condition, and half to the vowel-merge condition, but all completed a single-label block.

In each block, there were 16 training trials separated halfway through by a distractor trial, followed by 16 test trials. In the training trials, a cartoon creature moved onto the computer screen, paused in the center, and its label was stated twice, with a 1000ms delay between labelings (“Deev. ... Deev.”). After being labeled, the creature moved back off the screen. Then the next learning trial occurred. Learning trials occurred in a different random order for each participant. In single label conditions, 8 of the 16 trials presented one word and character, and the other 8 presented the other word and the other character. For merged word conditions, 4 trials were used for each specific word (e.g. fosh), summing to 8 across the merged phoneme for each competitor (4 fosh + 4 fush) as in the single label condition. At the 8-trial midpoint of training, a pair of distractor trials occurred to re-engage the participant. Each distractor trial consisted of an animal moving up and down across the screen with an entertaining sound effect.

The testing trials were a two-alternative forced choice task. Each test trial presented the two learned creatures stationary, side by side on the screen (position was counterbalanced within a test block) while one of the two creatures’ names was played once. The participant was asked to point to which creature was named, at which point the experimenter clicked the chosen creature. During test phases, children’s accuracy (points to pictures) was recorded along with visual fixations to pictures.

Results

We created generalized mixed-effects models with the package lme4 (Bates et al. 2014, R Core Team 2014). Throughout this study, when any two models are compared, finding that one accounts for significantly greater variance in a likelihood test (anova() function) between fixed effects describes the significance of each fixed effect. This comparison is achieved by creating two models: one model is made with all fixed and random effects present, and another identical model is created with the exception of a single fixed effect. The chi-square statistic and its associated p-value are reported for this comparison. Furthermore, Barr et al. (2013) suggest that likelihood ratio tests provide a more reliable estimate of statistical significance.

The first model assessed whether there was a general effect of label similarity. That is, is learning single labels easier for preschoolers than learning dual labels that are phonologically similar to each other? To test this, we modeled test accuracy in terms of the fixed effects of condition type (consonant merge or vowel merge) and merge type (variable labels or single labels). Fixed effects were sum-coded so patterns could be interpreted as they are in ANOVAs. Individual subject intercepts and word intercepts and slopes (both fixed effects and their interaction) were included as random effects. There was a main effect of merge type ($\beta=.30$, $SE=.12$, $Z=2.55$, $p=.01$). See Figure 13 for details.

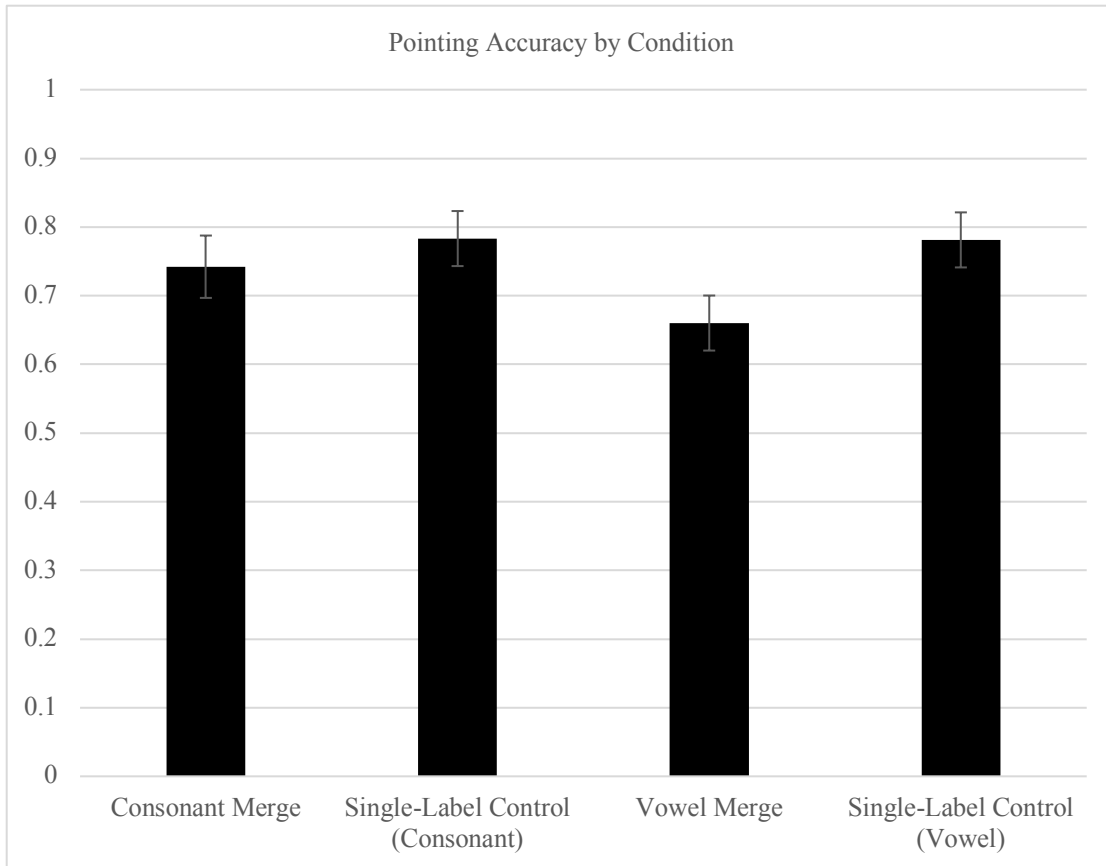


Figure 13. Accuracy in each merge condition. Collapsed across phoneme type, merging conditions were different from Single-label Control conditions, $p=.01$.

The model that removed the merge condition effect was significantly different from the larger model ($\chi^2=6.06$, $p=.01$), implying that accuracy is significantly diminished when merging sounds than when compared with single-label controls. The effects of condition type (consonant, vowel) were not significant, nor was the interaction. The lack of an effect of consonant vs. vowel variability stands in contrast to the work of Nazzi, who has found repeatedly that French-speaking children are less sensitive to vowel variability in distinguishing words.

Participants were also eye-tracked during the learning and testing portions of this experiment. To explore the participants' visual search patterns and assess whether participant looks corroborated their accuracy data, eye-tracking data from the testing portion of the

experiment were analyzed with mixed-effects models. Data were transformed following Barr (2008), and was done separately with participants and items as random effects to accrue multiple trials per data point. (Measuring looking proportion on a single trial results in a distribution that is neither binomial nor normal; aggregating across trials yields a distribution that is closer to normal.)

The starting model always contained maximal random effects structure. The model with the maximal random effect structure did not converge, so correlations between random effects were dropped, and then random effects themselves were dropped in order of least amount of variance explained until the model converged (Barr et al., 2013). In the first eye-tracking model we assessed the effect of segment condition (whether the participant learned a vowel or consonant merging vocabulary) and merge condition (whether the trial was testing a variable label, or a consistent [control] label) on the empirical-logit-transformation of looks to the target (Barr, 2008). Looks were analyzed in a 1-s (1000-ms) time window ranging from 200 to 1200 ms to encompass the full length of the single word. In Figure 14, an effect of merge type ($\chi^2=15.12$, $p<.001$) revealed more target looks overall when participants were in the control condition, mirroring the accuracy data.

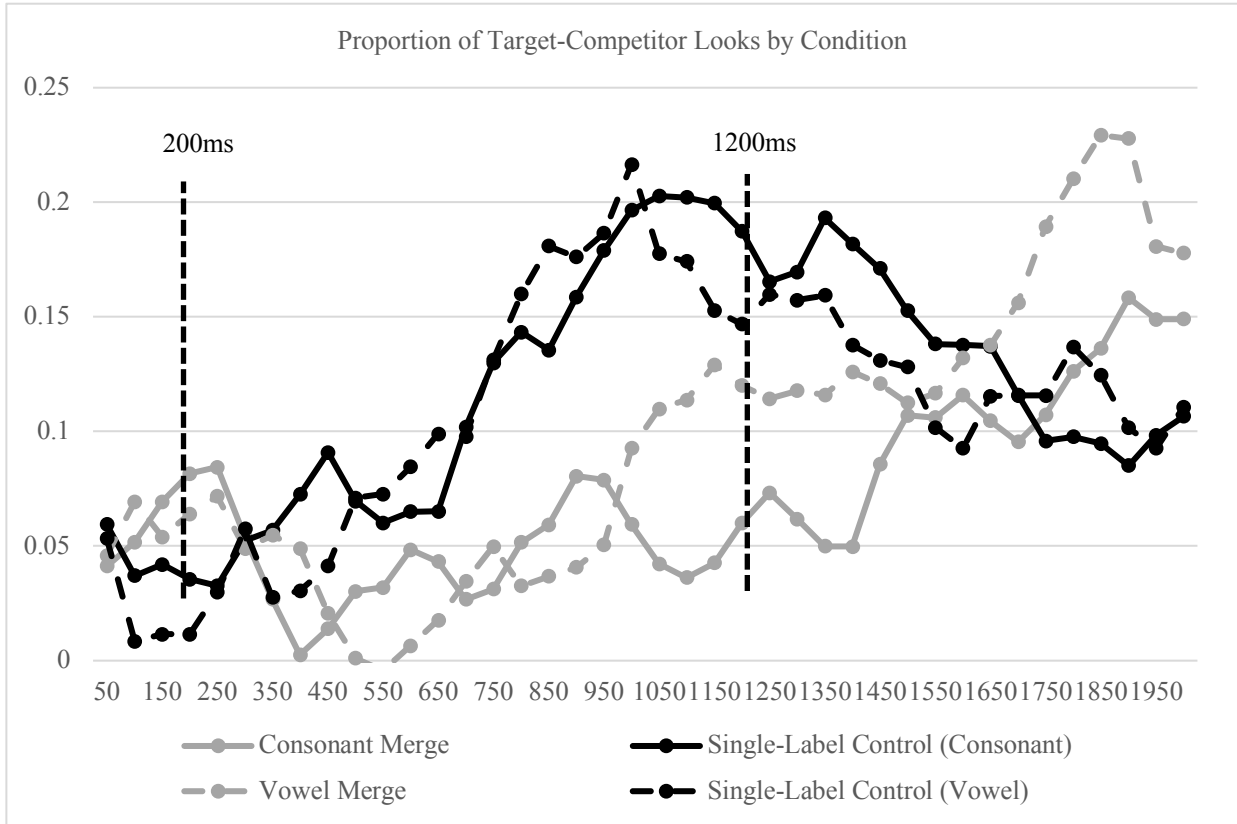


Figure 14. Correct-other proportion of looks in each merge condition. Collapsed across phonemes, merging labels received a smaller proportion of correct looks than Single-label Control conditions, $p < .001$.

In the second model, assessing the same effects with words as the random effects, the same effect was found ($\chi^2 = 12.63$, $p < .001$).

Discussion

We interpret the findings indicating some level of flexibility in children's word learning—while less accurate and slower to recognize words with variable phonemes, they are nonetheless well above chance. Yet due to the design of the experiment, the outcome of could also be explained by an inability to perceive the difference between minimal pairs in preschool-aged word learning. This would significantly reduce the strength of this data in terms of

explaining the developmental trajectory of phonemic flexibility in language learners. In order to demonstrate that this simplistic answer is not the case, we provided 3-to-5 year old subjects with a learning task where two creatures' labels were minimal pairs of each other. We predicted that they would be able to discern and learn the difference between these two labels and reliably identify the difference (recognize the creatures) when prompted, thus demonstrating our findings in this experiment represent a legitimate step in the word-learning developmental process.

This study also afforded another opportunity to examine differences in consonant vs. vowel flexibility. We tested each child on both a consonant minimal pair and a vowel minimal pair. If vowels are less integral to word identity, then children should learn the consonant minimal pair more accurately than the vowel minimal pair.

Experiment 5b

Method

Participants

We tested 32 monolingual English-speaking preschoolers ages 3-5 years (average: 4.46, SD: .82). Four additional participants were tested but dropped from the final dataset because they did not finish the experiment due to fussiness, 1 was dropped for technical issues, and 2 for excess exposure to languages other than English.

Stimuli

Visual stimuli are the same as those used in Study 5. Auditory stimuli were the 32 novel consonant-vowel-consonant (CVC) words used in Study 1, 4, and 5.

Procedure

The same procedure for training and testing was used as in Study 5. The minimal-onset condition asked children to learn names for two creatures that differed by onset consonant voicing (creature one: /div/, creature two: /tiv/). The minimal-vowel condition taught labels for two creatures that differed minimally in their vowels (creature one: /vaf/, creature two: /vΛf/). Each participant completed one block of minimal-onset word learning and one of minimal-vowel word learning, with block order counterbalanced across participants.

Results

Regression analyses were similar to those in Study 5. The model of accuracy assessed whether there was a general effect of label similarity. In this case, we wanted to see if it was difficult for preschool-aged children to distinguish minimal pairs in new words, which would weaken previous results. To test this, we modeled test accuracy in terms of the fixed effects of segment type (whether the participants were given minimal pairs by the onset consonant or by the vowel). Fixed effects were sum-coded so patterns could be interpreted as they are in ANOVAs. Individual subject and word intercepts and slopes (both fixed effects and their interaction) were included as random effects. There was no main effect, implying that accuracy is not significantly diminished when learning minimal pairs of either segment type. Children were above chance performance overall (59.86%), indicated by a significant intercept term ($\beta=.50$, $SE=.16$, $Z=3.16$, $p=.002$) and displayed in Figure 15.

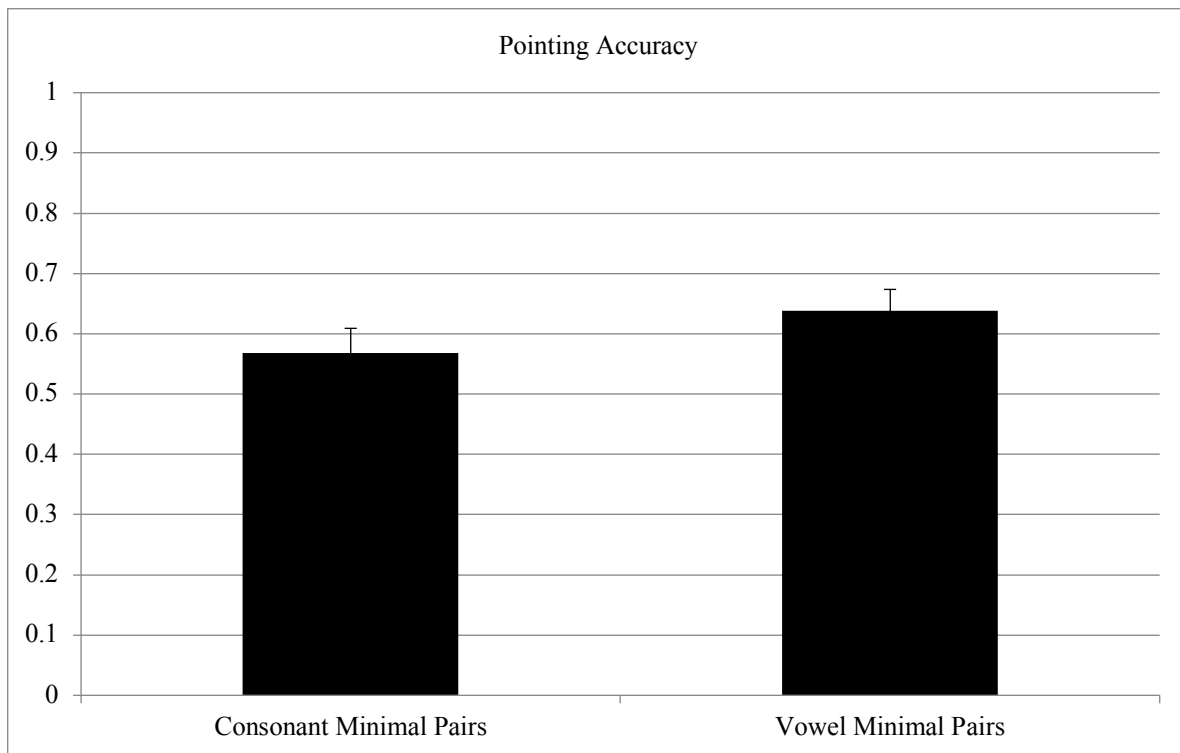


Figure 15. Pointing accuracy in each condition. Children performed above chance regardless of the minimal pair type, $p=.002$.

The starting model of eye tracking contained maximal random effects structure. Because the model with the maximal random effect structure did not converge, correlations between random effects were dropped, and then random effects were dropped in order of least amount of variance explained until the model converged (Barr et al., 2013). In the first eye-tracking model we assessed the effect of segment condition (whether the trial included a minimal pair of vowel sounds or consonant sounds) on the empirical-logit-transformation of looks to the target (Barr, 2008), based on subjects as the random effects. Looks were analyzed in a 1-s (1000-ms) time window ranging from 200 to 1200ms to encompass the full length of the single word. Mirroring the accuracy data, there were no effects of segment type for subjects or items. To assess whether children looked to targets at above-chance rates, we conducted a t-test collapsed across segment

condition comparing target looks to other-picture looks. This number did not exceed chance ($t = .35, p = .73$) (Figure 16).

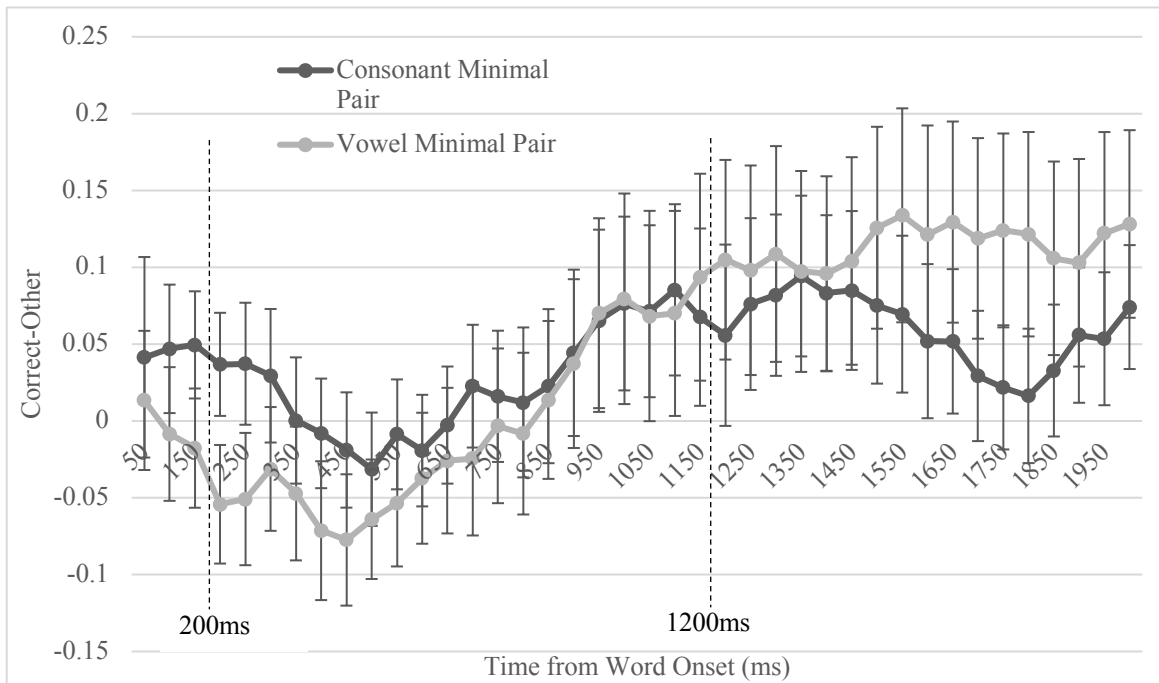


Figure 16. Correct-Other proportion of looks in each minimal pair condition. Neither condition showed a higher proportion of correct looks to target than chance, $p = .73$.

Discussion

One possible explanation for the results of the previous study (Experiment 5) was that participants at this age were simply incapable of distinguishing phonemes when presented as minimal pairs, and therefore their good performance on merged vocabularies was because they failed to distinguish differences between dual labels, rather than flexibility in word learning. (Of course, 3-to-5-year-olds failing to distinguish phonemes would be incongruous with any number of studies on even younger children.) However, the results of the current study demonstrate that 3-to-5-year-old children *are* capable of distinguishing minimal-pair words when required to learn them as distinct object labels. Interestingly, no effects of segment type (consonant or vowel)

were found for accuracy or for changes to patterns of eye tracking, providing evidence that this age group can discern between minimal pairs but that segment type is not yet meaningful.

Performance on this task, however, was not as robust as it could have been, hovering around 60% accuracy. This is somewhat surprising given evidence that much-younger 17-20-month-old infants can learn minimal pairs in other studies (Stager & Werker, 1997; Werker et al., 2002). Like some other aspects of learning language, such as regularization of irregular word forms, this might imply that 3-to-5-year-old children have entered a flexible moment in word learning and phonemic flexibility despite having the capacity to perceive meaningful differences between minimal pairs. It is clear that children of this age are capable of differentiating (and merging) minimal pairs when applied as object labels, but the poor performance on a task as seemingly simple as minimal pair distinction is an interesting twist on an otherwise straightforward result.

General Discussion

Now that it is clear the results from experiment 5 cannot be easily explained by an inability to perceive minimal pairs, we return specifically to the outcomes of the merged learning experiment. A key interest in this developmental study was to determine the possible time course of flexibility in language learners. We predicted that differences in segment perception throughout development might affect outcomes differently than adults for preschoolers learning phonemically-variable vocabularies. Different contributions of vowels and consonants in word-learning development has been proposed before (Nazzi, 2005), but the current study provides novel evidence that for developing speakers, when learning entirely novel vocabularies, children are equally sensitive to vowel versus consonant changes, which is different than what most

adults seem to experience as well as what previous studies have found for young children (though see Creel, 2012).

We have seen evidence in adults that vowels are easier to merge than consonants (Muench & Creel, 2013; Study 2), but previous work has indicated that perception of consonants is not adult-like until after age 12 (Hazan and Barrett, 2000). The results of this study demonstrate that merging consonants does not seem to cause preschool-aged learners substantial losses of learning accuracy, but neither does it decrease accuracy substantially for vowel-merging vocabularies. Without sufficient linguistic experience, it seems, phonemes in general are much more malleable than previously thought.

The results are inconsistent with studies of (e.g.) 16 and 20-month French-learning infants, who seem more likely to assume that words varying in their vowels are not different (Havy & Nazzi, 2009). This could represent differences between English and French in the relative weights of vowels and consonants, or differences in age groups, differences in tasks, or all three. The consistency of the Nazzi group's findings with French-learning children tends to suggest that native language (French vs. English) may be implicated.

The apparent difference between children's ease of recognizing mispronounced real words in Creel (2012) and mild difficulty in the current study could mean that 3 to 5-year-olds' perception of vowels is flexible enough to recognize similarity to known words, but that their learning mechanism is not yet quite flexible enough to allow for multiple versions of brand-new words.

Chapter 6: Conclusion

This dissertation sought to answer several related questions about the nature of phonemic representations. Specifically, I probed the extent to which phonemes are the atoms of word form identity, how the amount, type, and position of a phoneme affects its flexibility, how language background might influence all these relationships, and the developmental trajectory of this flexibility. This took place over the course of six experiments in five studies, four of those studies being aimed at the different aspects of flexibility in adult word learning, and one study with two experiments performed on 3-to-5-year-old children with the same goal. I now address each of these questions in turn, outlining what knowledge has been gained through the series of studies conducted.

The driving question behind this dissertation can be simplified to asking how crucial is stability in phonemes for learning words? The first experiment was designed to probe that question and set out to demonstrate that adults from multiple language backgrounds could learn vocabularies designed with flexible phonemes, with special attention paid to differing mutability between phoneme types. Broadly, this study made it clear that stability in phoneme presentation is not required to successfully learn words, and that flexible phonemes words can be used with no significant increase in difficulty to the learner. It also demonstrated a trend that continued throughout the studies where phonological competition in testing affected participants on a gradient: specifically, beginning in this first experiment and carrying throughout the dissertation, the importance of the onset consonant was reiterated by higher confusion when competitors shared this portion as well as simply an increased difficulty in responding accurately to the target when the competitor shared multiple phonemes with the target. Because we see this trend occur

throughout this document, our understanding of the importance of phoneme position during word learning is enhanced.

To answer a question raised by the first study, the second study probed the effect of different language backgrounds (English monolingual vs. Spanish-English bilingual) on learning with phoneme variability. Again, this study demonstrated that regardless of the subphonemic information present in the flexible vocabulary and its relationship to the language background of the learner, adults are just as capable of learning words with flexible phonemes as they are of learning stable phoneme vocabularies. The corroboration of data from the first study was crucial for pointing toward a larger generality underpinning language learning. Even the effect of differing degrees of competition on test accuracy was comparable between the two vocabularies and participant pools. However, the effect of language background was not as predicted: an increasingly large body of research has sought to demonstrate that there are advantages in word learning due to bilingualism. This would predict that in this dissertation, different language backgrounds would produce significantly different learning patterns or results. Multiple studies here demonstrate this was not the case. There may be specific advantages to being bilingual, but learning vocabularies with flexible phonemes is not one of them.

The third study explored the effect of syllable position versus phoneme type in the learnability of a flexible phoneme, and specifically set out to determine if vowels are more flexible than consonants, and whether coda consonants are more flexible than onset consonants. Importantly, even with a third new vocabulary, which contained flexibility in onset consonants and coda consonants, vs. onset consonants and vowels in the first two studies, the ability to learn the words as well as a vocabulary with stable phonemes was replicated. Although no main effect of segment position or type was found, the pattern of results did indicate that the position of a

phoneme is crucial to its effect on a learner when it is unstable within a vocabulary. The interactions of test competitor types and flexible phoneme position within a word imply that position is determining for recall difficulty within a phoneme type. Specifically, for test trials that contained competitors that matched the onset consonant, learners were less accurate than when test trials contained competitors that shared codas or vowels, regardless of the test condition they were trained in. This is to say that despite drawing the participants' attention to a specific aspect of a vocabulary, the most detrimental competitor to display was one with a shared onset consonant or multiple shared phonemes. It is also possible to demonstrate that coda positions cause less confusion in competition to a target than onset positions for consonants, implying that the position of the consonant is robustly affecting the learner.

The strength of adult word-learning in the face of flexible phonemes primed the fourth study, asking whether phoneme variability leads to more subtle difficulty in word recognition that is apparent in eye tracking. Mimicking the design of the first study and presenting the task to only monolingual English speakers, a comparable pattern of results to the first study emerged, specifically, that stability in phoneme presentation is not required to successfully learn words, and that flexible phonemes words can be used with no significant increase in difficulty to the learner. It also showed similar patterns of test confusion in the face of competitors sharing onsets or multiple phonemes, regardless of training condition. Overall, this study strongly corroborated the English monolingual data from the first study, lending confidence to the conclusions drawn by the first study via replication. Even regarding the eye-tracking data, which was designed to capture possible departures in learning that accuracy data alone could not perceive, the eye-tracking results directly supported the behavior captured by the accuracy data recorded.

With these results in hand, the next step was to determine how phoneme flexibility develops—do young children have difficulty learning words with variable phonemes, and are their consonant and vowel sensitivities different from adults? After making sure that children are in fact capable of distinguishing between—and learning—words with minimal pairs, the task of deciphering the extent of their ability to learn flexible phoneme vocabularies was explored. It was clear that the 3-to-5-year-old children were able to learn the flexible phoneme vocabularies, but their performance lagged behind their learning of stable phoneme vocabularies, implying that the capability of learning vocabularies with flexible phonemes was predicated on more extensive practice or language interaction than is available to children of that age.

This argument is supported by other models as well. Over development, there are some theoretical ideas about the architecture that would explain these effects similarly. Kleinschmidt and Jaeger proposed a system in children learning language that pays attention to situations as they relate to previously experienced ones and can adapt to novel situations (2015). In my studies, children were presented with novel speakers, but in situations that were not unfamiliar from others likely encountered by preschool-aged children in America. According to the *ideal adapter* framework proposed by Kleinschmidt and Jaeger, this would simultaneously be triggering known scenarios – associating cartoon characters with novel names as they move across a screen – and unknown talkers, provoking a need to adapt. The push and pull between these two systems, and indeed, the development of these systems within individual participants, could explain the above chance performance in learning while allowing for these child learners to be underperforming the adults.

The data also draw distinctions between our methods and results and those of other experiments attempting to answer related questions. As noted in background throughout this

work, the majority of research on phoneme flexibility has been done as a secondary question to the main experiment, and hasn't been able to make strong claims about specific flexibility in phonemes in the direct manner this work can. For instance, in terms of word learning, previous experiments that taught multiple labels did so in ways that differed from how this dissertation did. The few experiments teaching any sort of variable phonemes differed across place and manner of articulation (e.g. Muench and Creel, 2013), or treated vowel or consonant flexibility as secondary to the main experimental question in a way that didn't fully explore the capability of word learners for phoneme flexibility (e.g. Creel, Aslin, & Tanenhaus, 2006). The conclusions these studies arrive at, by-and-large that altering phonemes within labels for objects is significantly more difficult, likely occurs because of the methods used to test the learners. These data would most closely align with the dissimilar conditions given to participants in this work, and in fact our conclusions and theirs would largely align.

Where this work departs from previous efforts is in the proximity of phonemes to their variable labels, and the focus on simply that aspect of word learning. In fact, the only study to pair object labels similarly in vowels to this experiment found, like we did, that being presented with similar variable vowel labels was not detrimental to learning (Muench and Creel, 2013). Results from other work that most closely resembles aspect of this work often line up in terms of conclusions that can be drawn. The crucial difference in this dissertation is that it focused solely on the interactions of similar phonemes to word learning, largely different than the vast majority of related work, and so found results that were novel for the field in meaningful ways.

The results of each of these studies in regard to the novel treatment of flexibility of learners as measured by difficulty in learning variable phonemes specifically are represented in Table 4.

Table 4. A collection of results of each experiment in this work and their insight into the flexibility of word learners in specific flexible phoneme situations based on the comparison of merging, segment, and trial type effects.

Experiment	Merging Type Effect?
Phoneme Variability in Word Learning	Main effect
Native Language and Phoneme Variability in Word Learning	No main effect nor interaction
Syllable Position and Phonemic Variability Effects	No main effect, but an interaction with trial type
Phonemic Variability: A More Sensitive Measure	No main effect nor interaction
Developmental Effects on Phonemic Merging	Main effect

The data recorded in these studies across multiple language backgrounds and age groups have meaningful implications for the understanding of phonemes as building blocks of word representations. Over and over again, both adults and children, regardless of linguistic background, demonstrated a capability of learning new words despite comprehensive and varied forms of phonemic flexibility. However, this flexibility is not limitless, and these studies also demonstrated that it was possible to strongly inhibit learning with too much flexibility. It is vanishingly unlikely that subjects would be capable of managing a certain amount of flexibility in their word learning if the phoneme was a straightforward building block in word learning. Predictable flexibility in a vocabulary, regardless of phoneme type or position, allowed for the same accuracy as vocabularies with static phonemes, broadening our understanding of the capability of the human word-learning system and complicating the likely manner in which humans go about learning new words. It is clear that word learning is not a matter of placing

discrete sounds next to one another until a new combination is recognized; instead, we are clearly capable of much more flexible treatment of sound patterns when it comes to learning new words.

This data, however, did not demonstrate changes to learning capability based on language background. Despite ample evidence that a wide range of abilities are enhanced due to bilingualism, and even work that specifically cites word learning as a skill that is honed by being bilingual (Kaushanskaya & Marian, 2009), no evidence of increased capability in learning vocabularies with flexible phonemes was found throughout this work. It is possible that there was an insufficient range of languages tested to categorically deny any changes based on language background in word learning with flexible phonemes, but it is clear that Spanish-English bilinguals see no improvement in their accuracy on word learning compared to English monolinguals, despite differences in the number of vowels and consonants recognized as legal in those two languages. As discussed earlier, there is also the likelihood that the accuracy of certain participant groups is affected by the socioeconomic status differences likely present between the groups tested. This could lead to the conclusion that despite a native-language advantage being present, it is obscured by differing amounts of general language experience as predicted by differing socioeconomic standards. Although this work did not explicitly monitor socioeconomic status, it is likely that it had some effect on the data. Further study should explicitly note the status of the individuals to provide greater insight into the effects possibly caused by differing linguistic experiences due to socioeconomic status. Extended study could also improve chances for significant differences between languages by widening the gap between the compared languages, but the current study found no difference at all in performance.

The trajectory of this ability was explored in English speaking children and demonstrated that despite eventual similarities of word learning accuracy in the face of different phoneme type flexibility across language backgrounds this skill is not necessarily available to the word learner from birth. Children were still capable of learning vocabularies with flexible phonemes, but their accuracy on these flexible vocabularies was diminished relative to adults. This points to a learning period for sound flexibility and possibly even for word learning ability in English speaking children. It raises interesting questions as well about the aspects of words that children are paying attention to when learning words when compared with adults. Due to demonstrations that some flexibility is tolerated, but extreme flexibility makes it significantly harder to learn a vocabulary, the likely explanation for the difference in performance then shifts to more subtle aspects of words and learning in general. The current work does not illuminate what those may be, but only serves to rule out the most basic explanation of word learning in children.

The results from these studies add significant information to our understanding of word learning. This dissertation explores an area of language that the literature has hinted at, but as of yet remained unstudied. This document provides a solid empirical foundation for understanding the role of the phoneme in word learning across development, language background, phoneme type, and syllable location. The results provide useful evidence for multiple fields of study, including first and second language acquisition, language development, and bilingual vocabulary acquisition.

Bibliography

- BARTOLOTTI, J., & MARIAN, V. (2012). LANGUAGE LEARNING AND CONTROL IN MONOLINGUALS AND BILINGUALS. *COGNITIVE SCIENCE*, 36(6), 1129–47.
- BERGMAN, M. (1980). *AGING AND THE PERCEPTION OF SPEECH*. UNIVERSITY PARK PRESS.
- BEST, C. T. (1994). THE EMERGENCE OF NATIVE-LANGUAGE PHONOLOGICAL INFLUENCES IN INFANTS: A PERCEPTUAL ASSIMILATION MODEL. *THE DEVELOPMENT OF SPEECH PERCEPTION: THE TRANSITION FROM SPEECH SOUNDS TO SPOKEN WORDS*, 167, 167–224.
- BEST, C. T., McROBERTS, G. W., & GOODELL, E. (2001). DISCRIMINATION OF NON-NATIVE CONSONANT CONTRASTS VARYING IN PERCEPTUAL ASSIMILATION TO THE LISTENER'S NATIVE PHONOLOGICAL SYSTEM. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 109(2), 775–794.
- BIALYSTOK, E., CRAIK, F., GREEN, D., & GOLLAN, T. (2009). BILINGUAL MINDS. *PSYCHOLOGICAL SCIENCE*, 10(3), 89–129.
- BIALYSTOK, E., CRAIK, F. I. M., KLEIN, R., & VISWANATHAN, M. (2004). BILINGUALISM, AGING, AND COGNITIVE CONTROL: EVIDENCE FROM THE SIMON TASK. *PSYCHOLOGY AND AGING*, 19(2), 290–303.
- BOUCHON, C., FLOCCIA, C., FUX, T., ADDA-DECKER, M., & NAZZI, T. (2014). CALL ME ALIX, NOT ELIX: VOWELS ARE MORE IMPORTANT THAN CONSONANTS IN OWN NAME RECOGNITION AT 5 MONTHS. *DEVELOPMENTAL SCIENCE*, 1–12.
- CLOPPER, C. G., PISONI, D. B., & DE JONG, K. (2005). ACOUSTIC CHARACTERISTICS OF THE VOWEL SYSTEMS OF SIX REGIONAL VARIETIES OF AMERICAN ENGLISH. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 118(3), 1661–1676.
- CREEL, S. C. (2012). PHONOLOGICAL SIMILARITY AND MUTUAL EXCLUSIVITY: ON-LINE RECOGNITION OF ATYPICAL PRONUNCIATIONS IN 3--5-YEAR-OLDS. *DEVELOPMENTAL SCIENCE*, 15(5), 697–713.
- CREEL, S. C. (2014). PRESCHOOLERS' FLEXIBLE USE OF TALKER INFORMATION DURING WORD LEARNING. *JOURNAL OF MEMORY AND LANGUAGE*, 73(1), 81–98.
- CREEL, S. C., ASLIN, R. N., & TANENHAUS, M. K. (2006). ACQUIRING AN ARTIFICIAL LEXICON: SEGMENT TYPE AND ORDER INFORMATION IN EARLY LEXICAL ENTRIES. *JOURNAL OF MEMORY AND LANGUAGE*, 54(1), 1-19.
- CUTLER, A., SEBASTIÁN-GALLÉS, N., SOLER-VILAGELIU, O., & VAN OOIJEN, B. (2000). CONSTRAINTS OF VOWELS AND CONSONANTS ON LEXICAL SELECTION: CROSS-LINGUISTIC COMPARISONS. *MEMORY & COGNITION*, 28(5), 746-755.
- DELLE LUCHE, C., POLTROCK, S., GOSLIN, J., NEW, B., FLOCCIA, C., & NAZZI, T. (2014). DIFFERENTIAL PROCESSING OF CONSONANTS AND VOWELS IN THE AUDITORY MODALITY: A CROSS-LINGUISTIC STUDY. *JOURNAL OF MEMORY AND LANGUAGE*, 72, 1-15.

- DUNN, A.L. & FOX TREE, J.E. (2009). A QUICK, GRADIENT BILINGUAL DOMINANCE SCALE. *BILINGUALISM: LANGUAGE AND COGNITION*, 12(3), 273-289.
- DUPOUX, E., PEPERKAMP, S., & SEBASTIÁN-GALLÉS, N. (2010). LIMITS ON BILINGUALISM REVISITED: STRESS 'DEAFNESS' IN SIMULTANEOUS FRENCH-SPANISH BILINGUALS. *COGNITION*, 114(2), 266-275.
- ECKERT, P. (2008). VARIATION AND THE INDEXICAL FIELD I. *JOURNAL OF SOCIOLINGUISTICS*, 12(4), 453-476.
- EUROPEAN COMMISSION SPECIAL EUROBAROMETER. (2006). EUROPEANS AND THEIR LANGUAGES.
- FLEGE, J. E. (1991). AGE OF LEARNING AFFECTS THE AUTHENTICITY OF VOICE-ONSET TIME (VOT) IN STOP CONSONANTS PRODUCED IN A SECOND LANGUAGE. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 89(1), 395-411.
- FLEGE, J. (1998). THE PHONETIC STUDY OF BILINGUALISM. *ILHA DO DESTERRO*, 35, 19-28.
- FLORENTINE, M. (1985, SEPTEMBER). NON-NATIVE LISTENERS' PERCEPTION OF AMERICAN-ENGLISH IN NOISE. IN *INTER-NOISE AND NOISE-CON CONGRESS AND CONFERENCE PROCEEDINGS (VOL. 1985, NO. 3, PP. 1021-1024)*. INSTITUTE OF NOISE CONTROL ENGINEERING.
- FLORENTINE, M. (1985). SPEECH PERCEPTION IN NOISE BY FLUENT, NON-NATIVE LISTENERS. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 77(S1), S106-S106.
- FLORENTINE, M., BUUS, S., SCHARF, B., & CANEVET, G. (1984). SPEECH RECEPTION THRESHOLDS IN NOISE FOR NATIVE AND NON-NATIVE LISTENERS. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 75(S1), S84-S84.
- GOLINKOFF, R. M., MERVIS, C. B., & HIRSH-PASEK, K. (1994). EARLY OBJECT LABELS: THE CASE FOR A DEVELOPMENTAL LEXICAL PRINCIPLES FRAMEWORK. *JOURNAL OF CHILD LANGUAGE*, 21(01), 125-155.
- GOLLAN, T. H., WEISSBERGER, G. H., RUNNQVIST, E., MONTOYA, R. I., & CERA, C. M. (2012). SELF-RATINGS OF SPOKEN LANGUAGE DOMINANCE: A MULTILINGUAL NAMING TEST (MINT) AND PRELIMINARY NORMS FOR YOUNG AND AGING SPANISH-ENGLISH BILINGUALS. *BILINGUALISM: LANGUAGE AND COGNITION*, 15(03), 594-615.
- GONZALES, K., & LOTTO, A. J. (2013). A BAFRI, UN PAFRI BILINGUALS' PSEUDOWORD IDENTIFICATIONS SUPPORT LANGUAGE-SPECIFIC PHONETIC SYSTEMS. *PSYCHOLOGICAL SCIENCE*, 24(11), 2135-2142.
- GROSJEAN, F. (1988). EXPLORING THE RECOGNITION OF GUEST WORDS IN BILINGUAL SPEECH. *LANGUAGE AND COGNITIVE PROCESSES*, 3(3), 233-274.
- HALBERDA, J. (2003). THE DEVELOPMENT OF A WORD-LEARNING STRATEGY. *COGNITION*, 87(1), B23-B34.
- HARRELL, J., & CARRASQUILLO, O. (2003). THE LATINO DISPARITY IN HEALTH COVERAGE. *THE JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION*. 289, 1167.
- HAVY, M., & NAZZI, T. (2009). BETTER PROCESSING OF CONSONANTAL OVER VOCALIC INFORMATION IN WORD LEARNING AT 16 MONTHS OF AGE. *INFANCY*, 14(4), 439-456.
- HOFF, E. (2003). THE SPECIFICITY OF ENVIRONMENTAL INFLUENCE: SOCIOECONOMIC STATUS AFFECTS EARLY VOCABULARY DEVELOPMENT VIA MATERNAL SPEECH. *CHILD DEVELOPMENT*, 74(5), 1368-1378.
- HOUSTON, D. M., & JUSZYK, P. W. (2000). THE ROLE OF TALKER-SPECIFIC INFORMATION IN WORD SEGMENTATION BY INFANTS. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: HUMAN PERCEPTION AND PERFORMANCE*, 26(5), 1570.

- JOHNSON, J. S., & NEWPORT, E. L. (1989). CRITICAL PERIOD EFFECTS IN SECOND LANGUAGE LEARNING: THE INFLUENCE OF MATURATIONAL STATE ON THE ACQUISITION OF ENGLISH AS A SECOND LANGUAGE. *COGNITIVE PSYCHOLOGY*, 21(1), 60-99.
- KAUSHANSKAYA, M., & MARIAN, V. (2009). THE BILINGUAL ADVANTAGE IN NOVEL WORD LEARNING. *PSYCHONOMIC BULLETIN & REVIEW*, 16(4), 705–10.
- KLEINSCHMIDT, D. F., & JAEGER, T. F. (2015). ROBUST SPEECH PERCEPTION: RECOGNIZE THE FAMILIAR, GENERALIZE TO THE SIMILAR, AND ADAPT TO THE NOVEL. *PSYCHOLOGICAL REVIEW*, 122(2), 148–203.
- KUHL, P. K., STEVENS, E., HAYASHI, A., DEGUCHI, T., KIRITANI, S., & IVERSON, P. (2006). INFANTS SHOW A FACILITATION EFFECT FOR NATIVE LANGUAGE PHONETIC PERCEPTION BETWEEN 6 AND 12 MONTHS. *DEVELOPMENTAL SCIENCE*, 9(2), F13-F21.
- KUHL, P. K. (2000). A NEW VIEW OF LANGUAGE ACQUISITION. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES*, 97(22), 11850-11857.
- LABOV, W., ASH, S., & BOBERG, C. (2006). *ATLAS OF NORTH AMERICAN ENGLISH: PHONOLOGY AND PHONETICS*. BERLIN: MOUTON DE GRUYTER.
- LADEFOGED, P. (2001). VOWELS AND CONSONANTS. *PHONETICA*, 58(3), 211-212.
- LIBERMAN, A. M., HARRIS, K. S., HOFFMAN, H. S., & GRIFFITH, B. C. (1957). THE DISCRIMINATION OF SPEECH SOUNDS WITHIN AND ACROSS PHONEME BOUNDARIES. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY*, 54(5), 358–368.
- MACMILLAN, N. A., GOLDBERG, R. F., & BRAIDA, L. D. (1988). RESOLUTION FOR SPEECH SOUNDS: BASIC SENSITIVITY AND CONTEXT MEMORY ON VOWEL AND CONSONANT CONTINUA. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 84(4), 1262-1280.
- MADDIESON, I. (1984). *PATTERNS OF SOUNDS*. CAMBRIDGE UNIVERSITY PRESS.
- MARKMAN, E. M., & WACHTEL, G. F. (1988). CHILDREN'S USE OF MUTUAL EXCLUSIVITY TO CONSTRAIN THE MEANINGS OF WORDS. *COGNITIVE PSYCHOLOGY*, 20(2), 121-157.
- MARSLEN-WILSON, W., & ZWITSERLOOD, P. (1989). ACCESSING SPOKEN WORDS: THE IMPORTANCE OF WORD ONSETS. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: HUMAN PERCEPTION AND PERFORMANCE*, 15(3), 576.
- MAYO, L. H., FLORENTINE, M., & BUUS, S. (1997). AGE OF SECOND-LANGUAGE ACQUISITION AND PERCEPTION OF SPEECH IN NOISE. *JOURNAL OF SPEECH, LANGUAGE, AND HEARING RESEARCH*, 40(3), 686-693.
- MCMURRAY, B., TANENHAUS, M. K., & ASLIN, R. N. (2002). GRADIENT EFFECTS OF WITHIN-CATEGORY PHONETIC VARIATION ON LEXICAL ACCESS. *COGNITION*, 86(2), 33–42.
- MCQUEEN, J. M., NORRIS, D., & CUTLER, A. (1994). COMPETITION IN SPOKEN WORD RECOGNITION: SPOTTING WORDS IN OTHER WORDS. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY, AND COGNITION*, 20(3), 621.
- MIYAWAKI, K., JENKINS, J. J., STRANGE, W., LIBERMAN, A. M., VERBRUGGE, R., & FUJIMURA, O. (1975). AN EFFECT OF LINGUISTIC EXPERIENCE: THE DISCRIMINATION OF [R] AND [L] BY NATIVE SPEAKERS OF JAPANESE AND ENGLISH. *PERCEPTION & PSYCHOPHYSICS*, 18(5), 331-340.

- MUENCH, K. L., & CREEL, S. C. (2013). GRADIENT PHONOLOGICAL INCONSISTENCY AFFECTS VOCABULARY LEARNING. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY, AND COGNITION*, 39(5), 1585–1600.
- NESPOR, M., PEÑA, M., & MEHLER, J. (2003). ON THE DIFFERENT ROLES OF VOWELS AND CONSONANTS IN SPEECH PROCESSING AND LANGUAGE ACQUISITION. *LINGUE E LINGUAGGIO*, 2(2), 203-230.
- PISONI, D. B., & TASH, J. (1974). REACTION TIMES TO COMPARISONS WITHIN AND ACROSS PHONETIC CATEGORIES. *PERCEPTION & PSYCHOPHYSICS*, 15(2), 285–290.
- POLTROCK, S., & NAZZI, T. (2015). CONSONANT/VOWEL ASYMMETRY IN EARLY WORD FORM RECOGNITION. *JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY*, 131, 135–148.
- NAZZI, T., & BERTONCINI, J. (2009). PHONETIC SPECIFICITY IN EARLY LEXICAL ACQUISITION: NEW EVIDENCE FROM CONSONANTS IN CODA POSITIONS. *LANGUAGE AND SPEECH*, 52(Pt 4), 463–480.
- NAZZI, T., & NEW, B. (2007). BEYOND STOP CONSONANTS: CONSONANTAL SPECIFICITY IN EARLY LEXICAL ACQUISITION. *COGNITIVE DEVELOPMENT*, 22(2), 271–279.
- QUAM, C., & CREEL, S. C. (2017). TONE ATTRITION IN MANDARIN SPEAKERS OF VARYING ENGLISH PROFICIENCY. *JOURNAL OF SPEECH, LANGUAGE, AND HEARING RESEARCH : JSLHR*, 60(2), 293–305.
- RADEAU, M., MORAIS, J., & SEGUI, J. (1995). PHONOLOGICAL PRIMING BETWEEN MONOSYLLABIC SPOKEN WORDS. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: HUMAN PERCEPTION AND PERFORMANCE*, 21(6), 1297.
- ROGERS, C. L., LISTER, J. J., FEBO, D. M., BESING, J. M., & ABRAMS, H. B. (2006). EFFECTS OF BILINGUALISM, NOISE, AND REVERBERATION ON SPEECH PERCEPTION BY LISTENERS WITH NORMAL HEARING. *APPLIED PSYCHOLINGUISTICS*, 27(03), 465-485.
- SEBASTIÁN-GALLÉS, N., ECHEVERRÍA, S., & BOSCH, L. (2005). THE INFLUENCE OF INITIAL EXPOSURE ON LEXICAL REPRESENTATION: COMPARING EARLY AND SIMULTANEOUS BILINGUALS. *JOURNAL OF MEMORY AND LANGUAGE*, 52(2), 240-255.
- SEBASTIÁN-GALLÉS, N., VERA-CONSTÁN, F., LARSSON, J. P., COSTA, A., & DECO, G. (2009). LEXICAL PLASTICITY IN EARLY BILINGUALS DOES NOT ALTER PHONEME CATEGORIES: II. EXPERIMENTAL EVIDENCE.
- SLOWIACZEK, L. M., MCQUEEN, J. M., SOLTANO, E. G., & LYNCH, M. (2000). PHONOLOGICAL REPRESENTATIONS IN PRELEXICAL SPEECH PROCESSING: EVIDENCE FROM FORM-BASED PRIMING. *JOURNAL OF MEMORY AND LANGUAGE*, 43(3), 530–560.
- SLOWIACZEK, L. M., & HAMBURGER, M. (1992). PRELEXICAL FACILITATION AND LEXICAL INTERFERENCE IN AUDITORY WORD RECOGNITION. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY, AND COGNITION*, 18(6), 1239.
- SPOLSKY, B., SIGURD, B., SATO, M., WALKER, E., & ARTERBURN, C. (1968). PRELIMINARY STUDIES IN THE DEVELOPMENT OF TECHNIQUES FOR TESTING OVERALL SECOND LANGUAGE PROFICIENCY. *LANGUAGE LEARNING*, 18(s3), 79-101.
- SWINGLEY, D., & ASLIN, R. N. (2000). SPOKEN WORD RECOGNITION AND LEXICAL REPRESENTATION IN VERY YOUNG CHILDREN. *COGNITION*, 76(2), 147–66.
- SWINGLEY, D., & ASLIN, R. N. (2002). LEXICAL NEIGHBORHOODS AND THE WORD-FORM REPRESENTATIONS OF 14-MONTH-OLDS. *PSYCHOLOGICAL SCIENCE*, 13(5), 480–4.

- TABOSI, P., BURANI, C., & SCOTT, D. (1995). WORD IDENTIFICATION IN FLUENT SPEECH. *JOURNAL OF MEMORY AND LANGUAGE*, 34(4), 440-467.
- TAKATA, Y., & NÁBĚLEK, A. K. (1990). ENGLISH CONSONANT RECOGNITION IN NOISE AND IN REVERBERATION BY JAPANESE AND AMERICAN LISTENERS. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 88(2), 663-666.
- TANENHAUS, M. K., & SPIVEY-KNOWLTON, M. J. (1995). INTEGRATION OF VISUAL AND LINGUISTIC INFORMATION IN SPOKEN LANGUAGE COMPREHENSION. *SCIENCE*, 268(5217), 1632.
- TORO, J. M., NESPOR, M., MEHLER, J., & BONATTI, L. L. (2008). FINDING WORDS AND RULES IN A SPEECH STREAM: FUNCTIONAL DIFFERENCES BETWEEN VOWELS AND CONSONANTS: RESEARCH ARTICLE. *PSYCHOLOGICAL SCIENCE*, 19(2), 137-144.
- WEISS, D., & DEMPSEY, J. J. (2008). PERFORMANCE OF BILINGUAL SPEAKERS ON THE ENGLISH AND SPANISH VERSIONS OF THE HEARING IN NOISE TEST (HINT). *JOURNAL OF THE AMERICAN ACADEMY OF AUDIOLOGY*, 19(1), 5-17.
- WERKER, J. F., & TEES, R. C. (1984). CROSS-LANGUAGE SPEECH PERCEPTION: EVIDENCE FOR PERCEPTUAL REORGANIZATION DURING THE FIRST YEAR OF LIFE. *INFANT BEHAVIOR AND DEVELOPMENT*, 7(1), 49-63.
- WERKER, J. F., FENNELL, C. T., CORCORAN, K.M., STAGER, C. L. (2002). INFANTS' ABILITY TO LEARN PHONETICALLY SIMILAR WORDS: EFFECTS OF AGE AND VOCABULARY SIZE. *INFANCY*, 3, 1-30.