NORTHWESTERN UNIVERSITY

Accent with an Accent: A Study of the Interaction between Stylistically and Prosodically Conditioned Phonetic Variation

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILMENT OF REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Linguistics

By

Shawn Foster

EVANSTON, ILLINOIS

June 2025

Abstract

Speakers often deploy stylistically distinct pronunciation variants in order to communicate situationally relevant social information, altering the fine-grained phonetic details of their speech when performing a specific social role or to advance a communicative goal. Some of the same phonetic dimensions implicated in stylistic variation information are also used in the phonetic marking of prosodic structure. This dissertation examines how speakers navigate the interaction between these two sources of phonetic variation in order to communicate prosodic information while simultaneously adapting their productions to their current social situation using these limited and overlapping phonetic dimensions.

This thesis investigates the interaction between the phonetic marking of socially driven stylistic variation and the phonetic correlates of prosodic prominence with a focus on the segmental phonetic dimensions F1 and F2. Experiment 1 is a production study that examines variable phonetic marking of prominence among vowels in Mainstream American English. This study contributes to our understanding of the effect of prominence on segmental phonetic features and provides empirical grounding for three laboratory-based production studies that follow, which make use of an artificial language-learning paradigm. This paradigm is used to investigate the interaction between prominence and social information while controlling for other factors known to influence acoustic variation in vowel production, such as word frequency and neighborhood density, while also minimizing the influence of speaker's preexisting social knowledge and biases.

Experiment 2 is an artificial language study that serves as a validation of the experimental method used for the remainder of this dissertation, establishing the effects of prominence on vowels in an artificial language, which are compared to the parallel findings of prominence effects for vowels in English words from Experiment 1. Experiment 3 uses the same artificial language materials from Experiment 2 with the added feature that some words were produced with phonetically different vowel

variants by different model speakers. The aim of Experiment 3 is to compare the effects of prominence on vowel production for two phonetic variants of a single phonological category.

Experiment 4 builds on Experiment 3 by identifying the vowel variants with the differing social identities of the model speakers, i.e., as stylistic variants. The aim of Experiment 4 is to directly compare the influence of stylistic variation and prosodic prominence on vowel production. This study uses a between-subjects design in which some participants were presented with information meant to socially bias them toward phonetic variants with more peripheral vowels (one of the speaker-specific stylistic variants), the phonetic characteristics of which align with and reinforce the expected effects of prominence, while other participants received information meant to bias them towards pronunciations of the same words with less peripheral vowels (the other speaker-specific, stylistic variant). In prominence-lending contexts, the less peripheral vowel variants were in competition with the expected peripheralizing effects of prominence. Contrary to expectations, speakers biased towards word forms with less peripheral vowels still peripheralized their vowel productions under prominence, but did so in a more restricted way than was observed for similar vowel categories in the previous experiments. Overall, these results provide evidence that the phonetic strategies by which speakers mark prosodic prominence are flexible and sensitive to speakers' situationally specific, socially oriented communicative needs.

Acknowledgements

It is not hyperbole to say that this dissertation would have been utterly impossible without the truly, unbelievably supportive community in which I have been lucky enough to find myself.

My thanks go out first to my advisor, Jennifer Cole, for the generosity she has shown me in spending six years helping me to become the scientist and person I am today. I'm grateful for every hour she's spent wisely, kindly, and very, very patiently guiding me through my career as a student and researcher. I could not have imagined a better mentor. I also thank the other members of my committee, Matt Goldrick and Annette D'Onofrio, for the time and effort they have spent in helping me develop the best dissertation I could.

Next, I would like to express my gratitude to the most important person in my life. My journey through this PhD program would not have been possible without the support of my beloved wife, Alexis Foster. The past six years of writing about science nearly every day have not prepared me to find the words to express my gratitude for having her in my life. After over a decade together, she's shown me a capacity for love and loyalty that is stronger than anything I've ever experienced. It's been over a decade, and that love and dedication hasn't wavered.

I would also like to thank my mother and stepfather, Terry and Derrick Murphy, for everything from driving a U-Haul across six states so we could move to Chicago to being a place to come home to in Georgia. Mom, you've been there for me and Lexi through some of our most difficult times, and I'll never be able to say thank you enough. And my sister, Stephanie, for being there to listen when I wanted to ignore my research for a little while and talk about cartoons for what was, in retrospect, a bit too long.

And of course, I must thank the many friends I've made among my fellow grads. To Dan and Nicole Mirea Turner: Thank you for your friendship, for many nights spent watching movies and occasionally MMA, and for having me and Lexi at your wedding (sorry we couldn't make it in person! I recall there was a bit of a bug going around). To Thomas Sostarics, for his friendship and, honestly, probably, more

things than I can remember. But at the top of the list are sitting through an extra presentation of my defense slides, helping me take (and spectacularly fail) a fitness test, teaching me more about R than I had learned on my own in my whole career, and introducing me to karaoke. To Jennifer Dibbern, my friend and cohort-mate, with whom I've shared many a discussion of life and linguistics over boba tea. To Jaime Benheim, for a seemingly inexhaustible supply of enlightening conversation, a lot of which happened during late night Uber rides. And to Anatha Latshaw, for being a good syntax friend. Thanks as well to Ziyun Chew, Chantal DeLeon, Emma Wilkinson, Rebekah Stanhope and Züheyra Tokaç, for being good friends, office neighbors, lab mates, and the best model talkers a laboratory phonologist could ask for. And last but very far from least, Summer Green, my oldest friend, for listening to my entire grad school journey from start to finish.

Dedication

To Lexi, who has earned it

Contents

List of Figures

List of Tables

Chapter 1: Introduction

Effective speech communication often depends on fine-grained phonetic details which may not be critical to the perception of the lexical content of an utterance, but which play an essential role in establishing and updating interlocutors' understanding of their current situation and of one another. This dissertation examines how speakers utilize such within-category phonetic variation to fulfill multiple communicative functions when those functions depend on limited phonetic resources. I ask how speakers adjust their productions to navigate the potential tradeoff between the phonetic encoding of discourse-level pragmatic information associated with prosodic prominence and stylistic phonetic variation linked to a speaker's social context or interactional role. The interaction between these two provides a unique example of two sources of phonetic variation independently communicating multiple forms of information along several of the same phonetic dimensions.

Discourse-level pragmatic meaning and social context are closely interconnected, but in principle involve separable dimensions of meaning. Focus, as an aspect of information structure, is intimately related to the notion of common ground: Information which interlocutors agree is shared between them (Krifka, 2008; Roberts, 2012). Work exploring the interaction between social context and pragmatic meaning has argued that social and interactional context critically shapes what information a speaker will or will not assume their listener already knows (Beltrama, 2020). Despite the role that social context plays in shaping pragmatic communication, the two are capable of varying orthogonally to some extent. Speakers are at least theoretically capable of prosodically marking focus on any given constituent in a sentence while adopting their speech to a number of social contexts. Likewise, a speaker may, in principle, stylistically alter their productions to communicate a number of social messages within the same prosodic or pragmatic context. However, while the forms of information communicated by prosodic and stylistic variation are at least potentially separable, the effects each exert on the phonetic signal are not. Multiple phonetic dimensions, including F0, F1, F2, and duration, are known to play a role in encoding both social and pragmatic information.

The simultaneous communication of social and pragmatic information thus presents a dilemma**.** The effects that both exert on their shared phonetic dimensions can potentially compete with or reinforce one another. A given phonetic value may be ideal for signaling prosodic prominence but disfavored on social grounds. For example, a speaker may intend to adopt a more casual speaking style, which generally favors phonetic hypoarticulation (Lindblom, 1990; Smiljanić & Bradlow, 2005). Their pragmatic goals, however, may include marking certain material as prominent, which has been argued to favor phonetic hyperarticulation (Cho, 2005; de Jong, 1995). This creates a situation in which the signaling of one form of information must be adjusted based on context defined by the other. Speakers may shift their productions in ways that best encode prominence, in ways that are best adapted to their current social situation, or in a way that represents a phonetic compromise between both functions. Conversely, situations where both favor the same output create a potential source of ambiguity. In such situations, it is not clear if or how speakers ensure that their listeners understand whether they intend to convey pragmatic information via prosodic prominence, alter their speech for socially motivated reasons, or both. Regardless of the specific strategies speakers adopt, the fact that social and pragmatic information make use of these shared dimensions means speakers are faced with the question of how to adapt their use of the phonetic correlates of one in the face of context defined by the other.

Furthermore, it is possible that speakers must not only understand how to prioritize one form of information over another, but when to prioritize each dimension. The precise phonetic forms that mark social and pragmatic information have each independently been argued to vary as a function of other aspects of context. That speakers are repeatedly shown to adopt different phonetic forms when speaking to different interlocutors or inhabiting different social roles (Bell, 1984; Coupland, 1980; Hay et al., 1999;

Podesva, 2007; Rickford & McNair-Knox, 1994) suggests that socially motivated variation is at least in part sensitive to a speaker's judgements of their communicative needs in specific contexts.

While the relative importance of communicating pragmatic information may be similarly subject to variation between communicative contexts, this does not necessarily guarantee speakers can ignore or downplay the phonetic correlates of prosodic prominence that signal this pragmatic information when they judge social information to be more critical. They may be unwilling or unable to do so because the phonetic and phonological aspects of prosodic prominence constrain the speaker to a small range of acceptable phonetic values. This interaction may further be sensitive to features such as the frequency or predictability of the prominent word (Cole et al., 2010), the phonological characteristics of the underlying material (Lee et al., 2006; Mo et al., 2009), or the precise phonetic effects in question: Which factor favors shifting along which phonetic dimension and in which direction. The extent to which each whether prominence or socially motivated variation is prioritized may thus depend in large part on the specific linguistic and communicative context.

The primary goals of this dissertation are to ask how the phonetic communication of social information varies between prosodic contexts, how the phonetic correlates of prosody which convey pragmatic information vary between social contexts, and how this interaction varies based on a speaker's communicative needs. I investigate these questions through a series of laboratory-based speech production experiments that make use of an artificial language learning paradigm. This choice of method was motivated by the need to precisely control for the phonetic, phonological, and lexical factors that may mitigate the relationship between social information and prominence while experimentally manipulating speakers' social and communicative goals. In all experiments, speakers were required to produce forms that were either prosodically "neutral," encoding English broad focus, or prosodically marked to encode corrective focus. The dissertation consists of two studies, each of which consists of two experiments.

In order to address my primary research questions involving the phonetic interaction between social and pragmatic information, it is first necessary to understand the effects of each factor in isolation. The overwhelming majority of sociolinguistic work on English has investigated the relationship between social context and intracategory phonetic variation in the first and second formants (F1 and F2) of vowels, corresponding to vowel height and backness, respectively. While prominence distinctions are described in terms of the same phonetic dimensions, the question of the specific phonetic alterations that characterize prominence in English are still a subject of debate. Previous work on the topic has proposed multiple theoretical accounts of how prominence relates to the underlying segmental material and empirical evidence provides mixed support for these theories. The first study of this dissertation thus examines the phonetic features that mark prosodic prominence independent of their relationship to social factors. Experiment 1, the first within this study, looks at how the phonetic correlates of prominence change depending on the prominent vowel in the context of existing English words. This is followed by Experiment 2, which looks at a smaller subset of vowels and seeks to validate whether the observed effects of prominence also arise in the artificial language context used in subsequent experiments. The findings of Study 1 suggest that vowels tend to phonetically peripheralize under prominence by shifting away from the center of the vowel space in ways that reflect their phonological and phonetic features, but that not all vowels peripheralize to the same degree or along the same dimensions.

The goal of the second study is to address the question of whether and how speaker vary their phonetic marking of prosodic prominence and use of stylistic variation in order to balance the communication of social and pragmatic information based on the demands of a given social context. In order to phonetically encode prominence in different ways depending on their social situation, speakers must be able maintain distinct stylistic variants under prominence. Experiment 3 therefore asks whether the same distinctions in the expression of prominence that Study 1 found to exist between phonological categories also exist between phonetic variants of the same category. To ask this question in an artificial language context, I introduced speakers to two "dialects" of the artificial language in which the same phonological vowel categories (i.e., vowel phonemes) are realized in distinct phonetic forms. For example, a word which in one dialect has a tense vowel [i], has a lax [ɪ] in the other dialect. These forms differed along the phonetic dimensions of F1 and F2 that are also shown in Study 1 to convey prominence distinctions in English. On each trial, participants were biased towards one or the other variant by manipulating the relative frequency and recency of exposure to each variant. Experiment 3 also manipulated prosodic prominence by eliciting words under different prominence conditions. Experiment 3 found that prosodically neutral productions showed a recency effect whereby a participant's production on each trial gradiently varied between the two dialectal forms (e.g., tense [i] vs. lax [ɪ]), conditioned by the variant that was most recently heard. However, this recency effect was significantly weakened or non-existent for prominent vowels, suggesting that speakers may distinguish fewer distinct phonetic variants in prosodically prominent conditions.

Finally, in Experiment 4, speakers were exposed to an identical set of productions as Experiment 3. Experiment 4 differed in that rather than simply being exposed to two "dialects," of the language, speakers were informed that one variant was the "native," form while the other was "non-native." Depending on which variant was identified as native, the expected phonetic effect of prominence could either shift production towards or away from that variant. In contrast to the findings from Experiment 3, Experiment 4 found that speakers were capable of signaling prominence differently according to which form was preferred by their social context. When the expected effects of prominence were incongruent with the effects of social information, speakers appeared to alter the way they communicated prominence, sometimes resulting in vowel categories shifting in the opposite direction than would be expected based on the findings from Experiment 3. In the opposite case, speakers were found to maximize the phonetic marking of social information in a way that resulted in prominence no longer being appreciably marked along the dimensions it shared with social information. I argue that these results suggest that the phonetic encoding of prominence is highly flexible and that speakers make creative use of this flexibility in order to serve their contextually specific communicative needs.

In the remainder of this section, I review previous literature on the phonetic encoding of social information and pragmatic focus. In doing so, I argue that there are multiple dimensions along which social information and prosodic prominence could interact, but that previous work leaves several empirical gaps that must be filled before this interaction can be examined. I then present an overview of previous work on the phonetic correlates of prosodic prominence. I discuss multiple competing theoretical explanations for these effects and argue that empirical evidence from previous work does not conclusively support any of these theories. Following this, I briefly discuss previous work on the use of phonetic variation to communicate social messages, highlighting work that argues for various ways in which social information may be preferentially communicated in some linguistic environments over others.

Prosodic Prominence

Prosodic prominence is usually defined as the marking of one phonological constituent as "stronger" than another. In English, prominence is defined phonologically at the word level in the form of lexical stress and at the phrasal level in terms of phrasal prominence or phrasal stress. Phrase-level prosodic prominence is used to signify contextually specific pragmatic information in English, particularly focus and givenness (Büring, 2016). Phonetically, prosodic prominence is marked along both suprasegmental and segmental dimensions (Arvaniti, 2020; Cole, 2015). The former include pitch, duration and intensity, while the latter include alterations to vowel formants, among other phonetic changes such as increases in VOT (Cole et al., 2007). Phonetic markers of prominence are typically described as serving to make the prominent material more phonetically "extreme" or as serving as a form of phonetic "enhancement" (Cho, 2005; Erickson, 2002). Prosodic prominence is usually associated with longer segment duration and greater articulatory effort which is in turn correlated with more salient acoustic values. However, "phonetic enhancement" is an umbrella term for multiple types of phonetic alterations along multiple phonetic dimensions. The precise nature of many of these alterations is an area of active investigation. For example, it is uncontroversial that prominent material is generally perceived as longer, louder, and marked by greater pitch excursions than non-prominent material. Perception of duration and pitch, however, depend on multiple changes to the phonetic signal. For example, duration can lengthen different segments to different degrees (Turk & Shattuck-Hufnagel, 2007). Pitch excursions can likewise be marked by changes to pitch height or the timing of this height relative to the underlying segment**.** Such cues can interact with one another and may carry different perceptual weights in different situations (Arvaniti et al., 2024).

This work focuses on the segmental correlates of prominence; specifically, its effects on the formants of English vowels. These dimensions are interesting from a sociolinguistic point of view because they are also strongly associated with the use of stylistic variation to communicate contextually relevant social information. However, both the empirical and theoretical characterizations of the effects of prominence on these dimensions are significantly less well understood than those on suprasegmental ones. Like suprasegmental features, segmental phonetic alterations are generally theorized to favor more "enhanced" productions. F1 x F2 space, however, differs from suprasegmental dimensions in a number of ways which make the definition of "enhancement" less straightforward.

Unlike suprasegmental dimensions such as pitch and duration, F1 x F2 space is two dimensional. Vowels can potentially shift along F1, F2, neither, or both. Furthermore, dimensions like duration and intensity can be considered "monovalent," in that prominence is only associated with movement in one direction along these dimensions. That is, prominence material is not expected to be shorter or quieter than non-prominent, all else being equal. F1 and F2, however, can both be described as bivalent. That is, if "enhanced" is equated with "peripheral," speakers can potentially achieve this by shifting their productions in any direction along either axis: Raising, lowering, backing, or fronting. The question of determining in which direction a vowel will shift is deeply tied to a unique feature of the vowel space as a vehicle for prominence: Its relation to segmental contrasts. It is true that suprasegmental dimensions are not completely orthogonal to such contrasts. For example, although neither lexical tone nor vowel length are contrastive in English, pitch and duration serve as secondary cues to consonant voicing (Hanson, 2009) and certain vowel contrasts (Fox et al., 1995) respectively. F1 and F2, however, are the primary acoustic correlates that differentiate vowel categories in English. It is therefore at least possible for prominence to affect such contrasts and the extent to which it strengthens or weakens them depends in part on whether and how prominence manifests differently between vowel categories.

Theoretical accounts provide competing hypotheses concerning the acoustic dimensions of vowels that are affected by prominence, the direction or valency of change along those dimensions, and the extent to which these effects differ between vowel categories. Two primary theoretical accounts of the effects of prominence on vowels have been proposed. The first is the "sonority expansion" hypothesis, which proposes that vowels shift under prominence in ways that enhance *syntagmatic* contrasts between vowels and adjacent consonants (de Jong et al., 1993; Steffman, 2021a) . The sonority expansion hypothesis predicts that prominence is phonetically expressed by all vowel categories phonetically lowering (that is, exhibiting higher F1 values). In other words, it describes the effect of prominence as unidimensional, monovalent, and insensitive to vowel identity. This contrasts with the "contrast enhancement" hypothesis, which instead proposes that vowels shift to maximize *paradigmatic* differences between vowel categories that may occur in the same position (de Jong, 1995; Erickson, 2002). This phonetically manifests as vowel space expansion, similar to what has been reported in clear speech (Ferguson & Kewley, 2007; Smiljanić & Bradlow, 2005). Unlike sonority expansion, it predicts that prominence involves both F1 and F2 and that it can involve shifting in either direction along both dimensions. It also predicts that the direction in which a vowel shift depends on its phonological characteristics: F1 lowers in high vowels and raises in low vowels while F2 raises in front vowels while lowering in back vowels.

As its name suggests, the rationale behind the sonority enhancement hypothesis is that increasing F1 of all vowels makes them more sonorous. More sonorous vowels are more “vowel like,” and therefore contrast more sharply with the surrounding consonants. Greater sonority has in turn been hypothesized to facilitate listeners’ identification of the prosodic structure of an utterance (Cole et al., 2007). Much of the empirical evidence cited for this theory comes from kinematic data. (Erickson, 1998) report that prominence (referred to as “emphasis”) is associated with consistent lowering of the jaw in stressed syllables and raising of the jaw in surrounding deemphasized syllables. Although de Jong (1995) ultimately argues for a hyperarticulation account of at least some phonetic results of prominence, he also reports articulatory results suggesting that /ʊ/ is articulated with a lower tongue position under sentential prominence. There have also been some examples of acoustic evidence for this theory, including perception- and production-based studies. Lee et al. (2006) found that speakers produce /ɪ/ with raised F1 under prominence. Steffman (2021b) also finds that listeners perceive /ɛ/ to be more prominent when produced with a higher F1 value.

The studies cited above establish many key findings related to prominence effects on vowels, but they do not provide conclusive findings for the behavior of vowels across the vowel space. First, de Jong and colleagues (1993) argue that F1 may be more strongly associated with lowering of the tongue than lowering of the jaw and that the lowering of one articulator may not necessarily imply the lowering of the other. Productions studies have also tended to examine relatively small samples of speakers, leaving open the possibility that interspeaker tendencies might be obscured by intraspeaker variation. Lee et al. (2006) examined four, while de Jong (1995) examined three, and both studies found intraspeaker differences in the articulatory and acoustic properties associated with prominence. Furthermore, de Jong examined only a single vowel category /ʊ/. While /ʊ/ is a high back vowel and would therefore be expected to lower only under the sonority expansion account, it is a lax vowel, which leaves open the question of whether tense vowels exhibit similar sonority expansion effects. /ʊ/ is also notable for its extremely low functional load in English (Wedel et al., 2013). It may thus not be fully representative of the behavior of the vowel system as a whole. More critically, de Jong also reports that prominence was associated with lower F2 of /ʊ/. Lee et al. (2006) similarly find that F1 raising is most consistently seen in lax vowels. Both de Jong and Lee et al. suggest that their results could be explained if sonority expansion *and* hyperarticulation both play active roles in determining the acoustic effects of prominence.

This addresses one potential issue raised by the sonority expansion hypothesis: that raising F1 for all vowels regardless of their phonological features could potentially weaken vowel contrasts that depend on height. The contrast enhancement hypothesis avoids this problem by suggesting that all vowels peripheralize in accordance with their phonological specification. If the relative influence of both factors can vary between speakers, this may explain results such as Cho (2005) by which high vowels showed greater intra-speaker variability along the F1 dimension in relation to prominence. High vowels in this study were shown to phonetically raise *or* lower under prominence and that the pattern differed between the four speakers examined. High vowels in this study were also shown to consistently front under prominence, as predicted by the contrast enhancement hypothesis. Similar results were found in Lee et al., (2006) and Mo et al. (2009), which report that vowels tend to peripheralize in the front/back dimension. Lee et al. (2006) suggests that effects of hyperarticulation manifest primarily along the front/back dimension. Movement in the height dimension of the vowel space is essentially constrained to increase sonority only for non-high vowels.

Each strategy, sonority expansion and contrast enhancement, is compatible with multiple phonetic strategies. This problem is particularly pronounced in the case of contrast enhancement. There are multiple ways this enhancement could be defined, each of which would have different implications for how prominence will phonetically manifest. Essentially, enhancement can be defined vowel *intrinsically,* as a process of making a vowel phonetically distinct from other examples of that same vowel, or vowel *extrinsically*, as minimizing confusability from other vowel categories. In the case of the vowel intrinsic account, specific predictions about the extent to which a vowel will shift will depend critically on the distribution of phonetic values typical for that category: vowels will shift away from their "average" phonetic values. This on its own, however, does not make predictions about along which *direction* a vowel will shift or to what *extent*. In contrast enhancement accounts, the question of direction is usually predicted to depend on the phonological specification of the vowel. However, movement in accordance with a vowel's phonological features does not necessarily entail that it will be rendered more phonetically distinct from its neighbors. A vowel extrinsic account would predict that phonetic shifting under prominence will be constrained so as not to jeopardize contrasts between phonetically similar phonological categories. For example, in the case of lax vowels, such as /ɪ/ and /ɛ/, phonetic fronting stands to potentially decrease the distinction between these vowels and their corresponding tense counterparts, /i/ and /e/. Finally, the need to maintain phonological distinctions may serve as a constraint on *degree* of shifting, by ensuring that speakers do not peripheralize to an extent that would endanger segmental contrasts.

To summarize the above discussion, previous experimental studies of the phonetic effects of prominence have examined either small sets of speakers or analyzed only a limited subset of vowels in certain regions of the vowel space. While studies like Lee et al. (2006) and Mo et al. (2009) examined larger regions of the vowel space, they were constrained in their use of observational or corpus-based rather than experimental methods. In both studies, prominence was defined broadly without reference to the pragmatic or lexical context in which it was observed. This raises the possibility that the phonetic effects of prominence observed in each study were rooted in or influenced by heterogenous linguistic and extralinguistic factors. The first study of this dissertation is designed to answer the questions of to what degree, in what direction, and along what dimensions vowels shift under prominence by addressing these limitations. To do so, I employ an experimentally controlled production paradigm in which speakers produce words in prominent and non-prominent contexts that have been controlled for lexical and syntactic context.

Artificial Language Learning

The ideal situation for testing the phonetic interaction between prominence and social signaling is one in which two socially differentiated variants differ along a phonetic dimension that is also implicated in the signaling of prosodic prominence and can both be reliably elicited in a laboratory context. Ideally, the relationship between each variant and the social context in which it is observed would be exactly comparable. As an example, consider an experiment that examines how speakers mark prosodic prominence on two variants of the same phonological category: A monophthongized "Southern" version of the /aɪ/ vowel and more diphthongal "Northern" variant. In order to understand how the phonetic marking of prominence changes depending on both the variant a speaker selects and on the communicative demands of their situation, the relationship between variants and social contexts would need to be fully crossed. That is, the Southern Variant would be examined in a "Southern Preferred" context and a "Southern Dispreferred" context, likewise for the Northern variant. Importantly, the extent to which each form was preferred or dispreferred by a situation would need to be comparable. A "Southern" variant should be preferred in the "Southern Preferred" context to the same degree as the Northern variant is in the "Northern Preferred" context.

This example illustrates some difficulties that could arise when examining how prominence manifests differently on socially differentiated variants. It is difficult to devise two situations in which two different variants are preferred to the same degree, i.e., in which the "Southern Preferred" context favored the "Southern" variant exactly as much as the "Northern Preferred" variant favored the "Northern" variant. This would require a way to quantify how much a specific variant is preferred by a specific situation. While this may possible, it would require the experimenter to design two situations that favored different variants but were otherwise comparable along all dimensions that may impact socially motivated phonetic variation. This is because speakers' preexisting social knowledge of a given variant may cause them to associate a variety of social features with people who produce that variant that extend beyond simply being from the Southern United States, such as being uneducated or "laid back," through a process called second order indexicality by which linguistic forms come to be associated with meanings beyond their original contexts (Silverstein, 2003). This poses a problem for empirical investigation of such variants in a laboratory setting, as it becomes difficult to ensure that speakers will reliably produce a variant when its various social meanings may encourage them to shift their productions towards or away from that variant in potentially complicated ways.

Artificial language-based studies provide a way to partially circumvent the issue of speakers preexisting knowledge of the association between linguistic forms and their social meanings. Previous work has found that language users are able to make novel connections between linguistic forms and social categories over the course of an experiment. Samara et al. (2017) found that when speakers are exposed to socially conditioned variation in their input in an artificial language, they are capable of systematically reproducing that variation in their output. Similar findings have also been demonstrated in perceptual tasks using artificial social categories in natural language contexts (Docherty et al., 2013). Outside of perceptual work, language users have been shown to be sensitive enough to socially conditioned variation in linguistic input in artificial language to reproduce it in their own output. Work by Roberts and colleagues has used a paradigm in which participants interact with fictional "aliens" in order to experimentally investigate questions of how socially mediated variation spreads in different social and linguistic contexts (Lai et al., 2020), such as the ways in which variation can depend on the relative salience of a linguistic form (Li & Roberts, 2023) or the alienability of the trait with which it is associated (Sneller & Roberts, 2018).

Experiments like those described in Samara et al (2017) and the work of Roberts and colleagues examine socially conditioned variation at a categorical level. The experiments of this dissertation further extends an artificial language paradigm to investigate continuous, phonetic variation in speech production. An artificial language paradigm has the potential to be useful in this area for a number of reasons. The primary such advantage is that minimally relying on speakers' preexisting knowledge of socially motivated variation may be particularly critical when speakers are being asked to produce their own speech. As discussed above, language users often have ideologies surrounding particular social variants beyond the broad demographic characteristics that those variants may be associated with. Speakers may therefore be less willing to produce variants that they associate with negative qualities. There are findings to suggest, for example, that speakers' productions are less likely to converge to highly negatively stereotyped forms (Clopper & Dossey, 2020). The use of an artificial language also allows multiple phonetic variants to be associated with the exact same social role, thereby minimizing the problem of being unable to ensure that each variant is precisely equally suited to its "preferred" context. Outside of its advantage in minimizing language users' associations between phonetic forms and existing variants, this paradigm allows precise control over properties known to affect some of the same phonetic dimensions as prosodic prominence in ways that can appear very similar to prosodic prominence. These include distributional properties of the word, such as frequency and predictability (Aylett & Turk, 2004; Baker & Bradlow, 2009; Gahl, 2008; Jurafsky et al., 2001), as well as lexical neighborhood (Munson & Solomon, 2004).

However, while this approach affords a high degree of experimenter control, it sacrifices a degree of ecological validity, as discussed in Lai et al. (2020). Essentially, the abstract and minimally interactive nature of the task may cause speakers to behave differently than they would in "real world" interactive scenarios. For example, it is possible that the behavior of words under prominence does depend critically on speakers' preexisting knowledge of the word, such as its frequency or predictability, and that the behavior observed in the artificial language represents a sort of "default" strategy that does not accurately correspond those typically used in the real world. A particular area of caution in a speech production context is that, while it may be possible for speakers to rapidly associate new social categories with novel segmental phonetic categories, it may be more difficult to account for social meanings inherent to prominence itself. Prominence, as a form of phonetic enhancement, could be linked to social meanings of its own, such as "educated," "feminine" or "gay" (Schilling, 2013).

Prosodic Structure and Socially Motivated Variation

The primary question this dissertation seeks to answer is how the production of stylistic phonetic variation to signal social information is prioritized over or constrained by the phonetic expression of prosodic prominence. In a very broad sense, the term "style" describes a type of intraspeaker variation that serves some communicative function in a specific situation. Different theoretical approaches vary in the specific functions style is thought to serve as well as in the mechanisms that give rise to that stylistic variation. It has been described as *reactive*, done mainly in response to external factors (Eckert, 2012; Labov, 1981; Schilling, 2013) such as a specific location or the presence of an interlocutor. More recent approaches have conceived of style as more *proactive*, serving as a tool for speakers to signal complex social messages (Eckert, 2012). Outside of the variationist tradition, social information has been theorized to play a more indirect role in shaping intraspeaker variation. For example, Sumner et al. (2014) for example, argue that listener perceptions of "idealization" of a given variant affect the formation of representations of that variant.

Depending on the direction and degree of the specific phonetic variant associated with a given social function, stylistic variation can potentially endanger or strengthen phonological contrasts: moving some categories closer together and others further apart. Returning to the example of monophthongization, monophthongizing the /aɪ/ vowel threatens makes it more phonetically similar to the /a/ vowel, leading to potential confusion between words like "bind" and "bond." By utilizing monophthongization stylistically, a speaker may run the risk of their productions of such minimal pairs becoming confused. Chain shifts provide an example of stylistic variation that moves phonologically contrastive forms further apart from one another. For example, the Northern Cities Shift (McCarthy, 2011) involves vowels rotating about the vowel space such that as /æ/ encroaches onto the phonetic space typically occupied by /ɛ/, /ɛ/ shifts away from its original location. By implementing their productions in accordance with such shifts, speakers can potentially maintain phonological distinctions despite stylistic variation.

The question of to what extent the signaling of social information takes priority over the communication of a phonological contrast depends on how much stylistic variation a given category or structure permits and to what extent those stylistic alterations potentially obscure a contrast. The extent to which a phonological environment can be said to constrain the signaling of social information is the extent to which that environment limits a speaker's selection of stylistic variants. In the specific case of the interaction between social information and prosodic prominence, the contrast that social information stands to obscure or enhance is that between prominent and non-prominent material. For the communication of social information to be prioritized, the phonetic features that characterize prosodic prominence must allow further variation based on social context. This can take the form of a weakened phonetic expression of prominence, such that a prominent constituent may be less phonetically distinct from a non-prominent one, or it may simply be an *altered* expression of prominence¸ such that the prominence distinction is maintained within the socially conditioned style by increasing the phonetic distinction between prominent and non-prominent constituents along one or more dimensions that are not recruited as correlates of the stylistic variant.

Conversely, the extent to which prominence constrains the use of phonetic variation for social purposes is the extent to which prosodic factors condition a speaker to be unwilling or unable to select some stylistically differentiated forms. A series of studies by Clopper and colleagues provides potential examples of what such constraints may look like in the segmental domain. These studies examine speakers from different regional backgrounds within the United States and ask in which linguistic contexts they produce forms more typical of their "local dialect," and in which do they produce more "standard" forms (Clopper et al., 2017; Clopper & Pierrehumbert, 2008; Clopper & Tamati, 2014). These studies argue that more "dialectal" forms are associated with more "predictable" or "easy" contexts. Words with many confusable phonological neighbors are more likely to be produced in their "standard" form than those with fewer neighbors. Moreover, words spoken in "careful," speech style are more standard than those in "casual" style and words that are less predictable from preceding material are more standard than those that are more predictable.

Prosodically prominent environments are similar to "difficult," "careful," or "unpredictable," environments both in that prominence is associated with "less predictable" information and in that prominence is argued to result in phonetic hyperarticulation. Consequently, results which suggest that more "standard" forms are associated with hyperarticulating environments (Clopper et al., 2017; Clopper & Pierrehumbert, 2008; Clopper & Tamati, 2014) lead us to predict that speakers would "prioritize" the signaling of prosodic structure. However, Clopper and colleagues caution that although the various sources of phonetic reduction that they examine may surface in similar ways, they may originate from different mechanisms. In particular, they state that variation due to speaking style is more readily conceived of as an overtly listener-oriented process than is variation due to factors such as lexical competition (Clopper et al., 2017). It is therefore possible that shifts towards more "standardized" forms observed in careful speech style is on some level reflective of speaker belief that their listener will have an easier time recovering their intended message if they produce a more standardized form. The extent to which speakers restrict their use of socially meaningful variation in prominent contexts may thus depend on to what extent the phonetic variation associated with prominence is listener-oriented in nature, as opposed to, for example, primarily a phonetic reflex of phonological structure.

Jacewicz et al (2006), in contrast, present evidence that regional shifts are *more* advanced in prominent conditions. The discrepancy between these findings and those of Clopper and colleagues may be partially attributable to differences in the phonetic characteristics of the investigated sound changes. The sound changes investigated by Clopper and colleagues are, for the most part, congruent with the phonetic alterations associated with easy contexts. For example, Clopper and Pierrehumbert (2008) found that monophthongization was observed in Southern Speech in predictable contexts, but the same was true to a lesser degree of the Midlands speakers of the same study. Likewise, /u/ fronting is simultaneously a phonetic alteration associated with predictable context and an ongoing phonetic shift throughout most of North America. The sound changes investigated by Jacewicz et al., however, included the phonetic lowering of /ɛ/ in the American South, which exerts phonetic pressure in the same direction as effects related to prominence (e.g., sonority expansion for lax vowels).

With such conflicting findings in mind, a few further caveats need to be taken into account when applying these results to the case of prosodic prominence. The first is simply that prosodic structure may not be prioritized to the same extent that phonological contrasts encoding lexical information are. Secondly, prominence is also robustly signaled by suprasegmental dimensions, such as F0, intensity and duration, which with the exception of duration, do not play a primary role in lexical or phonological contrasts in English. Furthermore, speakers in the above-mentioned experiments were not explicitly encouraged towards or away from a given form by their communicative goals. This leaves open the possibility that speakers can effectively "resist," the expected effect of prominence when doing so is preferred by the social context.

Regardless of the extent to which social information is prioritized over the information communicated by prominence (related to referential meaning (Büring, 2016) and/or semantic intensity (Sandberg, 2024), it is possible for prominence to constrain socially based stylistic variation either as an aspect of syntagmatic phonological context or as a consistent source of interference at the phonetic level. This distinction stems from two features that have been suggested to structure the relationship between a phonetic form and the social meaning it communicates. On one hand, the social message a given production communicates has been argued to be “analog,” (Eckert & Labov, 2017) or “incremental”, with gradient phonetic changes signaling corresponding gradient changes in social meaning. Under such a view, “variants,” can be described as endpoints along a phonetic continuum rather than a set of discrete alternatives conveying categorically distinct social messages.

On the other hand, the relationship between phonetic form and social meaning is also mediated by social and phonological context. In terms of social context, the same phonetic production may be interpreted differently depending other aspects of the social context, such as observable social qualities of the speaker or the interactional context. In terms of linguistic context, as Eckert and Labov argue, two phonetically identical segments can be assigned different meanings depending on the phonological context in which they occur. The case of monophthongization in the American South provides a concrete example of this. While monophthongization of /aɪ/ in open syllables and before voiced consonants is associated broadly with “The South,” monophthongization before voiceless consonants is more specifically associated with the Inland South or Appalachian region. Importantly, monophthongization before voiceless consonants can be considered more phonetically marked than in pre-voiced contexts, as speakers of English from outside the Southern United States have been shown to exhibit less dynamic vowels in this context (Pycha & Dahan, 2016). Unconditioned merger can therefore be said to represent a “more advanced,” instance of the sound change. Speakers would therefore not be expected to monophthongize in voiceless contexts without doing so in voiced consonants.

The fact that the social interpretation of a phonetic production depends on both its phonological context and its phonetic details has consequences for predictions about the interaction between prominence and social information. If prominence is treated strictly as a source of phonetic variation, then when two stylistically differentiated variants are separated along a phonetic dimension that is also implicated in signaling prominence, it would be impossible for speakers to mark prominence in a "socially neutral," way. All movement along that dimension would move speakers closer to or further from one variant and thus potentially affect the social message it sends. The fact that prominence has been proposed to serve as a mechanism to draw listener attention to a particular aspect of an utterance (Kristensen et al., 2013) means that prominent material may serve as a "high risk, high reward," site for the communication of social information. However, prominence is also part of a segment's phonological context. This may allow the messages signaled by prominent material to differ from those that would normally be associated with a production in non-prominent contexts.

In summary, the fact that prosodic prominence and social stylistic variation are communicated along the same dimensions implies that they must interact in some way. This dissertation asks, "in what way," and "to what end," do they do so. What strategies do speakers employ to encode these two dimensions of sociopragmatic meaning into overlapping phonetic dimensions? The remaining chapters of this dissertation answer these questions by first elucidating the means by which speakers encode prominence more generally. I then go on to investigate how the interaction between prominence and stylistic variation dynamically reflects a speaker's social situation, the social goals they pursue and the social roles that they inhabit.

Chapter 2: Phonetic Effects of Prominence Across the English Vowel Space

As stated in the previous chapter, prior work suggests that social information and prosodic prominence are communicated along some of the same phonetic dimensions. Specifically, both are associated with variation in F1 and F2. It is thus possible for their effects to reinforce one another by favoring the same phonetic outcome or detract from one another by favoring different outcomes. The phonetic outcomes associated with social information are argued to be highly flexible, with speakers altering their productions in different ways depending on the specific social context in which they find themselves and the specific social goal they intend to pursue. The phonetic effects of prominence are poorly understood in comparison. The segmental phonetic outcomes associated with prominence are subject to debate and have been argued to serve multiple, sometimes conflicting communicative functions. It is also unknown to what extent these segmental correlates of prominence are flexible, being expressed differently based on communicative context and goals.

Understanding the interaction between social information and prosodic prominence thus requires an empirical understanding of the expected effects of prosodic prominence independent of social context. One of the major debates surrounding prosodic prominence is its relationship to the phonetic and phonological characteristics of the underlying segmental material with which it is associated. Study 1 is designed to answer the question of how prosodic prominence is expressed phonetically in different phonological and lexical environments. It consists of two production experiments designed to elicit prosodic prominence through its association with corrective focus. Experiment 1 asks what sorts of phonetic alterations characterize prominence in different segmental contexts by comparing its effects across thirteen English vowel categories. Experiment 2 similarly relies on the relationship between prominence and corrective focus but investigates its effects on novel words in an artificial language context. This serves primarily as a proof of concept for the experiments that make up the remainder of this dissertation by establishing that the kinds of variation associated with prominence in English can be observed in the artificial language learning paradigm used by the experiments discussed in Chapter 3.

Experiment 1: Introduction

In Experiment 1, I seek to characterize the ways in which prosodic prominence shifts the production of various English vowel categories along the dimensions of height and backness, as measured through the first and second formants (F1 and F2). The hypothesized effects of prominence are based on findings from prior work that have posited that these segmental effects reflect two potential phonetic strategies: Sonority expansion and contrast enhancement. The expected effect of sonority expansion is to increase vowel sonority by raising F1 for all vowel categories and thereby enhancing the acoustic distinction between vowels and consonants. The contrast enhancement account predicts instead that vowels will shift under prominence so as to maximize paradigmatic enhancement between vowel categories. As discussed in the previous section, I posit that contrast enhancement can be further divided into vowel intrinsic and vowel extrinsic hypotheses. Unlike sonority expansion, both predict that the direction in which a vowel will shift depends on its phonological or phonetic characteristics. Front vowels should exhibit higher F2 under prominence and back vowels lower F2; High vowels should show lower F1 and low vowels raise F1. A vowel intrinsic account predicts that vowels shift so as to become more phonetically distinct from typical values of the same vowel. In its simplest form, this account predicts that vowels shift along both F1 and F2 and in doing so enhance the recoverability of their phonological features. The vowel extrinsic account predicts that vowels shift in ways that maximize perceptual distinctions between confusable vowel pairs. This predicts that vowels may fail to move in the direction associated with their phonological specification if doing so would bring them closer to a perceptually similar vowel. For example, lax vowels may fail to move under prominence if doing so would bring them closer to a confusable tense vowel (for the purpose of illustration, here ignoring other phonetic dimensions that may distinguish contrastive vowels).

Sonority expansion and contrast enhancement thus offer qualitatively different answers to the questions of which segmental dimensions are affected by prosodic prominence and whether those effects differ between vowel categories. It is also possible that the expression of prominence reflects a combination of both strategies. Contrast enhancement may, for example, only surface when it does not conflict with the expected effects of sonority expansion. This would predict that non-low vowels would only shift along F2, avoiding the F1 region of low vowels, while low vowels may shift along F1 (increasing their distance from non-low vowels) or along both dimensions. Similarly, Lee et al (2006) provides evidence that lax vowels may be more strongly affected by sonority expansion than tense vowels. Previous work does not, however, provide clear predictions to determine the degree to which either strategy for expressing prominence would be active for any given vowel, preventing the formulation of more detailed, quantitative predictions of the phonetic effects of prominence, such as the precise degree to which a particular vowel will shift along a particular dimension.

Understanding which, if any, of these prominence strategies is active in the phonetic expression of prominence requires an analysis of the entire vowel space. This experiment does so through eliciting prominent and non-prominent productions of thirteen vowel categories: {a, æ, ʌ, aʊ, aɪ, ɛ, e, ɪ, i, o, oɪ, ʊ, u}, which includes all vowels of Mainstream American English except /ɔ/ and rhotacized vowels. These were excluded due to the widespread merger of /ɔ/ and /a/ in North American English and to avoid phonetic interference from following rhotics. The question this experiment addresses is along what dimensions, in what directions, and to what degree each of these vowels shifts under prominence. I compare the acoustic consequences of prominence for all vowels in order to answer the additional question of how the effects of prominence reflect the phonetic and phonological properties of each vowel category.

I also present two supplementary hypotheses concerning the effect of prominence on F0 and vowel duration. Compared to its segmental correlates, it is relatively well accepted that prosodic prominence is associated with increases along both of these suprasegmental acoustic dimensions. Thus, investigating these correlates of prominence is meant to confirm that any segmental alterations we observe are due to the effect of prominence, rather than stemming from another source such as hyperarticulation due to clear speech. We also investigate the relationship between an individual vowel's segmental and suprasegmental variation under prominence by asking whether vowels that show smaller magnitude of variation along segmental dimensions the former show greater magnitude of variation along suprasegmental dimensions the latter.

Methods

Participants

Participants were recruited remotely via Prolific (N=40). Data from eleven were rejected due to audio quality issues or technical errors. All participants were compensated for their time in accordance with Northwestern's minimum wage guidelines. Of the twenty-nine participants whose data was analyzed, fifteen identified as male, thirteen as female, and one as non-binary. Mean age was 34.7 (s.d. = 7.7). All participants reported being raised monolingually in the United States. Two reported knowledge of other languages. Participants were classified into having grown up in one of four regions: Northwest (nine), South (thirteen), Northeast (nine) and West (four). This totals to more than twenty-nine because some participants reported multiple regions. Although speakers from the South constituted a plurality, no region provided a majority. This helps ensure that any observed effects of prominence are not specific to any specific regional dialect.

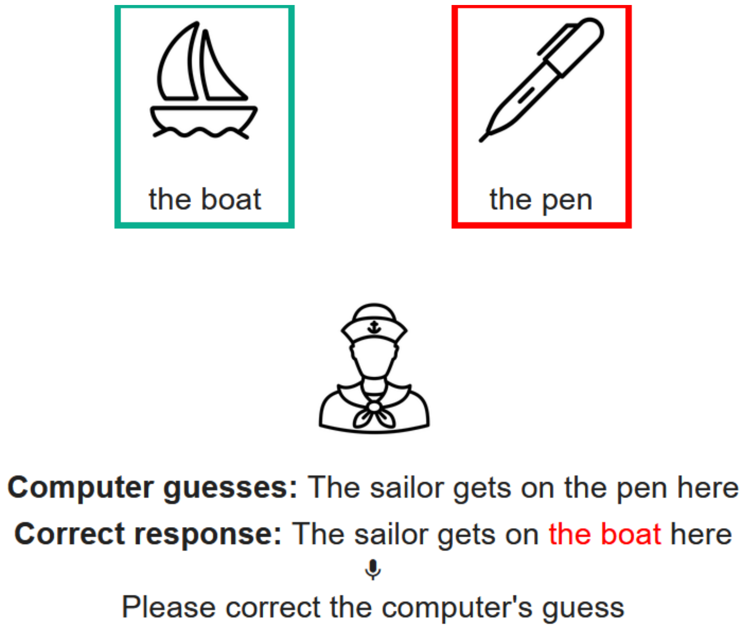
Stimuli

Visual stimuli consisted of simple black and white vector graphics taken from The Noun Project, a website providing stock photos and images. Figure 1 provides an example of a "prominent" trial. Each of the twenty-six trials contained three images for a total of seventy-eight images. All images appeared on only one trial. Each trial contained an orthographic prompt consisting of a single critical sentence. The sentence was always of the following form.

“The [subject noun] [verb + particle] the [critical word] here”

Where the subject was always a one or two syllable noun. The verb was always a single syllable and always paired with a single syllable particle (I.e., “picks up, “looks like,” “talk to.”) The phonological form of the verb phrase was kept constant to control for potential prosodic interactions between the verb and the critical object. Critical words were always one syllable nouns of the form “CVC” that were more closely semantically related to the subject noun than the distractor item was. The first and final consonants of each one syllable word were always obstruents. V was always one of the thirteen vowels examined by the study. Each vowel category appeared in two critical words. Whenever possible, one word contained the vowel in pre-voiced contexts and the other in pre-voiceless contexts. Due to the need for words to be relatively common, picturable nouns, this was not possible for /ʊ/, which appeared only before voiceless consonants (“bush, “book”) and /oɪ/, which appeared only before voiced consonants (“toys,” “boys,”). The word “here” was included at the end of every stimulus to control for effects of sentence final position on the production of creaky voice in English. A complete list of written stimuli is given in Appendix A.

Figure 1: Example of a "prominent" trial



Procedure

The experiment was conducted via a custom-made web application. After completing a consent form and background questionnaire, participants went through an instruction phase in which they were told they would be helping train a computer to “describe simple scenes.” They were told the machine learned by looking at pictures and listening to spoken phrases. Participants were then shown an example screen of a “scene,” which consisted of three black and white images arranged in a triangle. The two pictures at the top of the triangle were accompanied by orthographic representations of the pictured object. Under these images was a short English sentence containing a blank. The image at the bottom of the screen represented the subject of the sentence and the two objects on top represented two choices of words that could fill in the blank. Participants were told that the computer’s goal was to describe the scene in a “logical,” way. In the context of this experiment, a “logical way” meant that completing the sentence with the incorrect noun resulted in a sentence that described an impossible or highly unlikely situation, while inserting the other noun resulted in a possible or common outcome. The participants’ goal was to confirm or correct the computer's guess. Participants were told that the computer learned by listening to both "what they said," and "how they said it," and that it was important that they correct the computer's guess as though they were correcting a person.

During these instructions, participants were walked through two example trials. The first illustrated a computer providing a "correct," guess, which constituted a "non-prominent" trial. In this situation, the correct image would be highlighted in green, the blank would be filled in with the correct term, and the computer's guess would appear under the correct response. On correct trials, the word selected by the computer appeared in green. They were then shown an incorrect or "prominent" trial. On these trials, the correct image was highlighted in green while the incorrect was highlighted in red. The blank for the correct guess was filled with black text and the computer's guess appeared underneath in red. Participants were then shown another example, this time with accompanying audio from a model talker to indicate exactly when and how they should respond. The model talker produced one correct trial, produced with broad focus and one incorrect trial in which the object of the sentence was marked with corrective focus. Words in these examples, "horse," and "shirt," were chosen because no critical trial contained vowels in pre-rhotic contexts. This was done on the assumption that the distinct phonetic realization of these vowels compared to pre-obstruent contexts would minimize the likelihood of the model talker unintentionally influencing participants' productions. Participants then moved on to the first round of critical trials. These were identical in form to the first example they saw. Each of twenty-six critical words appeared twice: once in a "correct," trial and once in an "incorrect" trial. This resulted in a total of 52 trials per round. The order of trials was randomized between participants. After all trials had been completed, participants were asked to take a short (one minute) break before beginning the next round. This was identical to the first except that the trial order had been re-randomized. This process was repeated five more times for a total of six rounds. Median completion time was 47 minutes.

Analysis

Analyzed tokens consisted of recordings of critical words. Twenty-nine speakers producing a maximum of 312 productions each would have provided a maximum of 9048 total tokens. Tokens were excluded that were of poor audio quality, could not be force-aligned, or were more than 2 s.d. away from the mean of a speaker's entire vowel space. All vowels were plotted individually by-speaker and visually inspected before and after removal of outliers. Prior to removal of outliers, extreme outlying tokens were checked by hand to confirm that they represented technical errors or mistakes. 718 tokens were removed in this way, resulting in a total of 8330 tokens and a loss rate of 9.2%. Recordings were force-aligned using the Montreal Forced Aligner (McAuliffe et al., 2017). Acoustic measurements were extracted via a Praat script. Measurements of F0, F1, and F2 were taken at each 10% interval vowel portion of critical tokens.

The data was analyzed using five statistical models: Three examining all vowels and two examining only diphthongs. The first was a multivariate mixed-effects model predicting F1 and F2 at vowel midpoint. All 8330 usable tokens were included in this analysis. The second and third were univariate mixed-effects models predicting F0 at vowel midpoint and log vowel duration. Of the 8330 tokens included in the F1-F2 analysis, automatic pitch extraction through Praat failed to recover midpoint F0 for 411. These tokens were excluded from the F0 analysis, resulting in a total of 7919. All tokens included in the F1 – F2 analysis were also included in the log duration analysis. F0, F1 and F2 were scaled by speaker and reported in standard deviations from the mean of the speaker's vowel space or that speaker's mean F0 value. Each model included the interaction between trial type (a two-level categorical predictor with levels "prominent," and "non-prominent,") and phoneme (a thirteen-level categorical predictor) as fixed effects and speaker as a random effect with slopes for trial type and phoneme. The remaining two models examined only the diphthongs /o/, /e/, /aʊ/, /oɪ/, and /aɪ/. Like the initial multivariate model, each of these looked at F1 and F2. Unlike this model, F1 and F2 were measured at the 20% mark for one model and 80% for the other. Intercept and coefficients for all models included weakly informative normally distributed priors centered on zero. All models were implemented using the brms (Bürkner, 2017) interface to the Stan probabilistic programming language (Carpenter et al., 2017). Loosely informative student-t distributions were used as priors for standard deviation on all models. Interactions and posterior means were calculated using the emmeans package.

Results

Figure 2: Mean normalized F1 and F2 values of monophthongs in Experiment 1

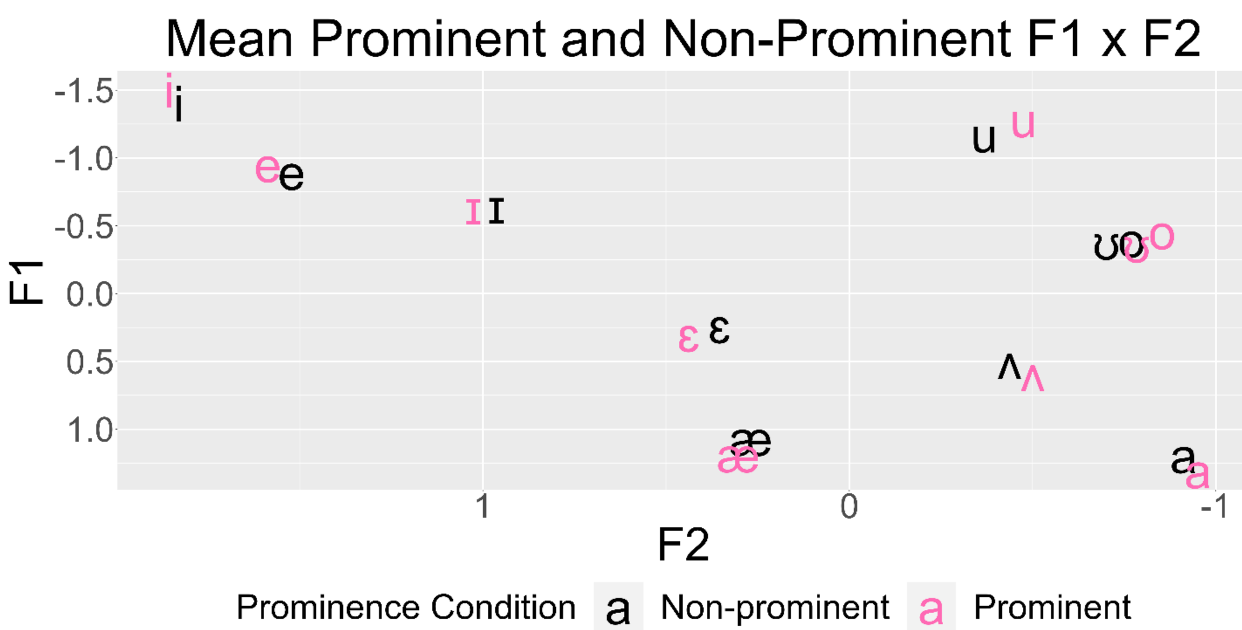
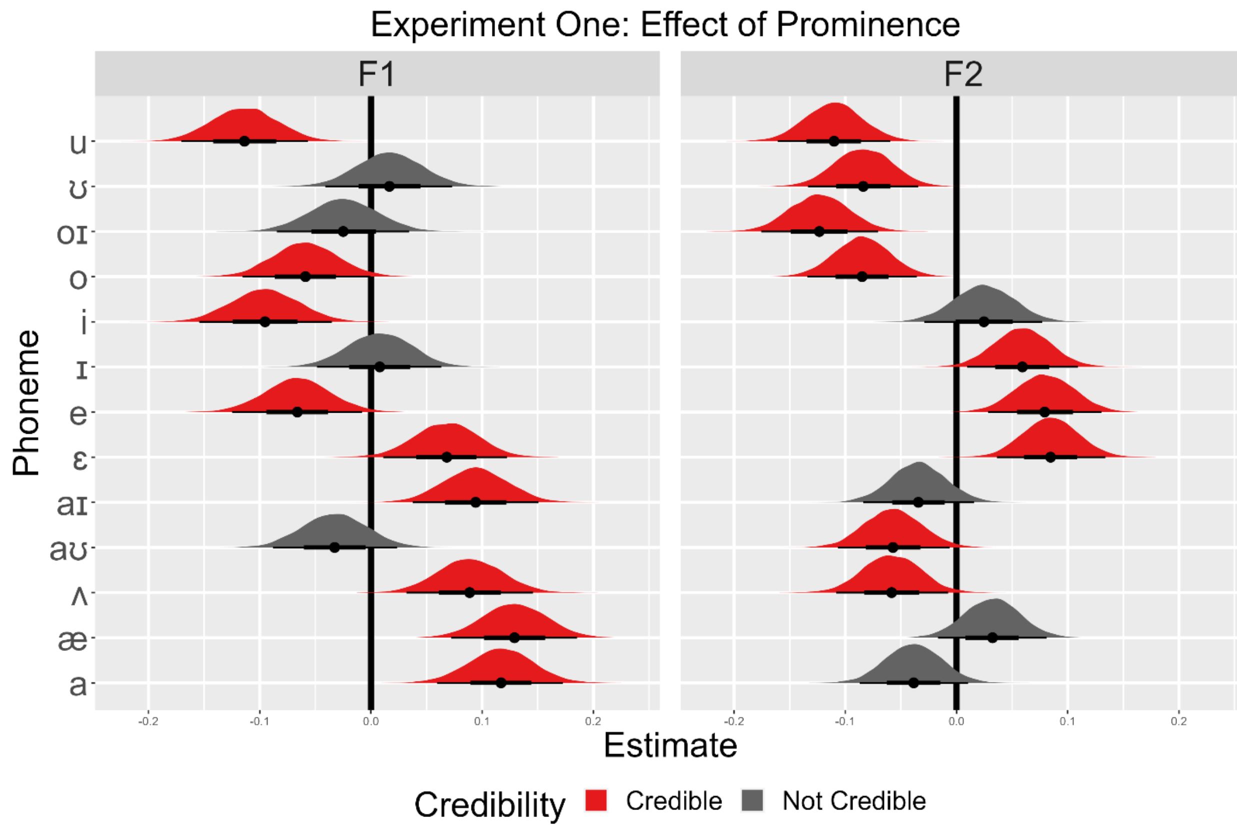


Figure 2 depicts empirical average midpoint F1 and F2 for all monophthongs in prominent and non-prominent conditions normalized by speaker. As Figure 2 suggests, peripheralization was observed along both F1 and F2 across the vowel space. Figure 3 illustrates results of the multivariate mixed-effects model predicting F1 and F2. Specifically, it shows the distributions of the model-estimated differences in posterior means between prominent and non-prominent conditions for each vowel. Distributions further to the left indicate that prominent vowels were found to be further back (F2) or higher (F1) than non-prominent vowels. Similarly, distributions further to the right indicate that prominent vowels are on average further front (higher F2) and lower (higher F1) than non-prominent vowels. Credible effects of focal prominence are shown in red. A credible effect can be thought of as similar to a significant effect in frequentist models. Distributions whose 95% credible intervals cross zero (illustrated by the vertical black line in Figures 3) are taken to represent non-credible effects.

Figures 3 illustrates that all vowels exhibited an effect of prominence along at least one formant dimension. Nine of the thirteen vowels examined showed movement along F1 under prominence. High vowels /i/ ($\beta = -0.1$, 95% CI = [-0.154, -0.035]) and /u/ ($\beta = -0.11$, 95% CI = [-0.17, -0.057]), mid tense vowels /e/ ($\beta = -0.07$, 95% CI = [-0.125, -0.008]) and /o/ ($\beta = -0.06$, 95% CI = [-0.115, -0.001]), and mid diphthong /oɪ/ ($\beta = -0.02$, 95% CI = [-0.085, 0.034]) raised in the height dimension, with credibly lower F1 under prominence. Lax mid vowels /ɛ/ ($\beta = 0.07$, 95% CI = [0.011, 0.122]) and /ʌ/ ($\beta = 0.09$, 95% CI = [0.032, 0.146]), low vowels /æ/ ($\beta = 0.13$, 95% CI = [0.072, 0.185]) and /a/ ($\beta = 0.12$, 95% CI = [0.059, 0.173]), and diphthong /aɪ/ ($\beta = 0.09$, 95% CI = [0.038, 0.151]) all lowered in the height dimension, with credibly raised F1. The two high lax vowels /ɪ/ and /ʊ/ as well as the remaining diphthongs /aʊ/ and /oɪ/ did not credibly move along F1 at vowel midpoint, despite the fact that visual inspection of Figure 3 may suggest that /ʊ/ was phonetically lower under prominence.

Figure 3: Effect of prominence on all vowels in Experiment 1



Nine vowels also moved credibly along F2. The set of vowel categories that failed to move along this dimension was disjoint from the set that did not move along F1. Five vowels moved along both dimensions. These include /u/ (F2: β = -0.11, 95% CI = [-0.161, -0.06]), /o/ (F2: β = -0.08, 95% CI = [-0.134, -0.036]) and /ʌ/ (F2: β = -0.06, 95% CI = [-0.108, -0.007]), which backed under prominence, as well as /e/ (F2: β = 0.08, 95% CI = [0.028, 0.13]) and /ɛ/ (F2: β = 0.08, 95% CI = [0.036, 0.134]) which fronted. The remaining vowels, /i/, /aɪ/, /æ/ and /a/ did not credibly move along F2 despite doing so along F1.

Table 1: Posterior Estimates for Effect of Trial Type on F1 of All Vowels

Phoneme	β	95% CI
a	**0.117**	**0.059, 0.173**
æ	**0.129**	**0.072, 0.185**
ʌ	**0.089**	**0.032, 0.146**
aʊ	-0.033	-0.088, 0.024
aɪ	**0.094**	**0.038, 0.151**
ɛ	**0.068**	**0.011, 0.122**
e	**-0.066**	**-0.125, -0.008**
ɪ	0.008	-0.048, 0.063
i	**-0.095**	**-0.154, -0.035**
o	**-0.059**	**-0.115, -0.001**
oɪ	-0.025	-0.085, 0.034
ʊ	0.017	-0.041, 0.073
u	**-0.114**	**-0.17, -0.057**

Table 2: Posterior Estimates for Effect of Trial Type on F2 of all Vowels

Phoneme	β	95% CI
a	-0.039	-0.087, 0.01
æ	0.032	-0.017, 0.081
ʌ	**-0.058**	**-0.108, -0.007**
aʊ	**-0.057**	**-0.107, -0.006**
aɪ	-0.034	-0.084, 0.016
ɛ	**0.085**	**0.036, 0.134**
e	**0.079**	**0.028, 0.13**
ɪ	**0.059**	**0.01, 0.109**
i	0.025	-0.029, 0.077
o	**-0.085**	**-0.134, -0.036**
oɪ	**-0.123**	**-0.175, -0.071**
ʊ	**-0.084**	**-0.134, -0.034**
u	**-0.11**	**-0.161, -0.06**

Table 3: Posterior Estimates for Effect of Trial Type on F1 of Diphthongs

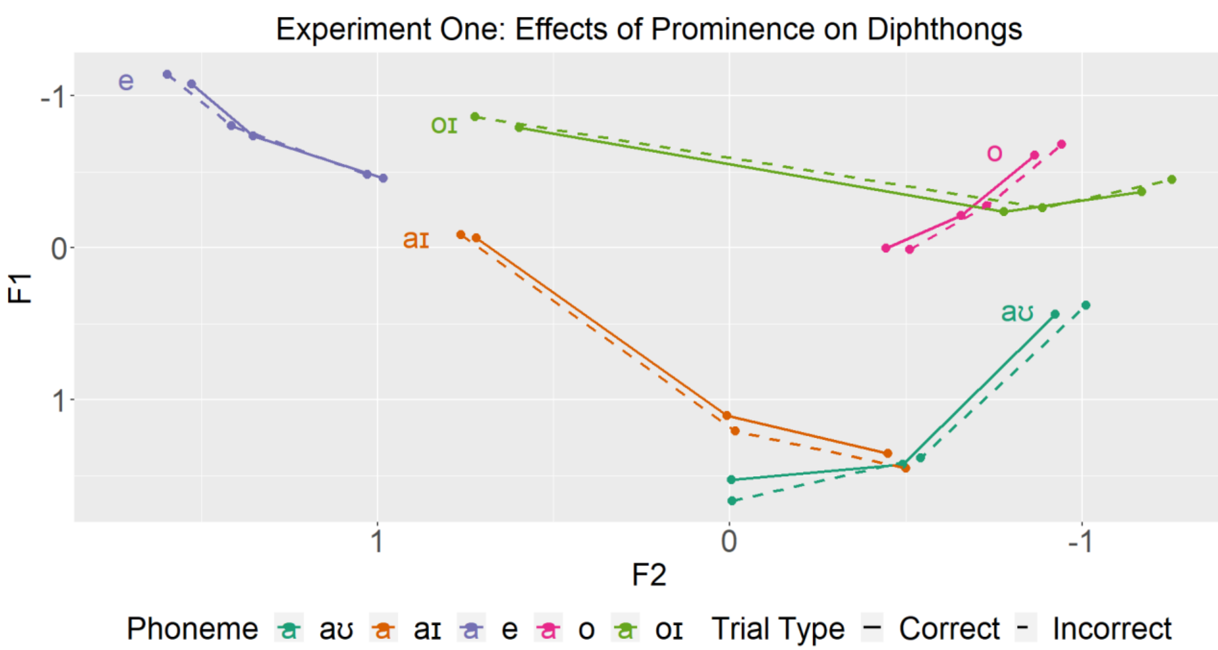
	20% Point	
Phoneme	**β**	**95% CI**
aʊ	0.126	**0.052, 0.198**
aɪ	0.082	**0.007, 0.157**
e	-0.025	-0.101, 0.051
o	0.003	-0.069, 0.076
oɪ	-0.082	**-0.159, -0.007**
	50% Point	
Phoneme	**β**	**95% CI**
aʊ	-0.033	-0.088, 0.024
aɪ	0.094	**0.038, 0.151**
e	-0.066	**-0.125, -0.008**
o	-0.059	**-0.115, -0.001**
oɪ	-0.025	-0.085, 0.034
	80% Point	
Phoneme	**β**	**95% CI**
aʊ	-0.047	-0.117, 0.024
aɪ	-0.014	-0.084, 0.056
e	-0.063	-0.134, 0.009
o	-0.073	**-0.143, -0.004**
oɪ	*-0.06*	*-0.133, 0.012*

Table 4: Posterior Estimates for Effect of Trial Type on F2 of Diphthongs

	20% Point	
Phoneme	**β**	**95% CI**
aʊ	-0.004	-0.08, 0.072
aɪ	-0.065	-0.14, 0.009
e	0.06	-0.018, 0.137
o	-0.072	-0.146, 0.003
oɪ	-0.107	**-0.188, -0.028**
	50% Point	
Phoneme	**β**	**95% CI**
aʊ	-0.057	**-0.107, -0.006**
aɪ	-0.034	-0.084, 0.016
e	0.079	**0.028, 0.13**
o	-0.085	**-0.134, -0.036**
oɪ	-0.123	**-0.175, -0.071**
	80% Point	
Phoneme	**β**	**95% CI**
aʊ	-0.094	**-0.168, -0.023**
aɪ	0.042	-0.031, 0.115
e	0.079	**0.005, 0.154**
o	-0.092	**-0.165, -0.018**
oɪ	0.142	**0.065, 0.219**

Figure 4 shows the average normalized F1 and F2 values for diphthongs at the 20%, 50%, and 80% duration points. Vowel labels mark the 80% points. The vowels /aʊ/ and /oɪ/, under focal prominence, were found to credibly move along F1 at their 20% marks despite not doing so at their midpoints. F1 of /aʊ/ was higher under prominence at this point and that of /oɪ/ was lower. The reverse was found for /e/ and /o/, which raised credibly at their mid points but did not credibly move at the 20% marks. /aɪ/ moved at both timepoints. Only /o/ was found to shift along F1 at the 80% mark, doing so in the same direction as its midpoint. No vowel moved along F2 at 20% except /oɪ/, which phonetically backed. All except /aɪ/ credibly moved at the 80% mark. Three of these, /o/, /e/ and /aʊ/, moved in the same directions in which they moved at their midpoint. The vowel /oɪ/ was unique phonetically backing at the 20% and 50% while fronting at its 80% mark.

Figure 4: Effect of trial type on diphthongs



There was no credible main effect of trial type on F0 nor did any interaction between trial type and vowel result in a credible difference between prominent and non-prominent trial types along this dimension. A credible effect of duration was found in the expected direction such that prominent productions are longer than non-prominent tokens (β = 0.13, 95% CI = [0.068, 0.196]). Interactions indicate this effect holds for every vowel. Differences between the duration of prominent and non-prominent productions were similar for all vowels.

Table 5: Posterior Estimates for Effect of Trial Type on Duration of all Vowels

Parameter	β	95% CI
a	0.11	**0.035, 0.184**
æ	0.105	**0.029, 0.181**
ʌ	0.123	**0.048, 0.198**
aʊ	0.124	**0.05, 0.2**
aɪ	0.117	**0.042, 0.193**
ɛ	0.132	**0.056, 0.206**
e	0.104	**0.028, 0.179**
ɪ	0.162	**0.085, 0.236**
i	0.153	**0.077, 0.23**
o	0.163	**0.087, 0.238**
oɪ	0.137	**0.061, 0.213**
ʊ	0.118	**0.043, 0.192**
u	0.173	**0.096, 0.248**

Discussion

The primary goal of this experiment is to empirically establish the effect of prosodic prominence on the segmental phonetic features of vowels via its relationship to focus. Prominence was hypothesized to reflect one of three strategies: sonority expansion, vowel intrinsic contrast enhancement, and vowel extrinsic contrast enhancement**.** Sonority expansion predicted that vowels would consistently lower in the height dimension under focus while the contrast enhancement predicted that vowels would peripheralize in the direction that would either enhance the perceptibility of a vowel's phonological features (vowel intrinsic hypothesis) or better distinguish between vowel categories (vowel extrinsic hypothesis). In other words, **a**s discussed in Chapter 1, the effects of sonority enhancement are monovalent, unidirectional, and insensitive to vowel identity while those of contrast enhancement are bivalent, bidirectional, and sensitive to vowel identity. The current results, I argue, while not completely compatible with the predictions of any hypothesized strategy, come out more clearly in favor of a contrast enhancement hypothesis.

Some vowels were found to phonetically lower under prominence, as predicted by sonority expansion. However, three vowels which lowered, /a/, /æ/, and /aɪ/ at midpoint, were phonologically low vowels. They would therefore have been predicted to do so under any of the three strategies, as lowering was also predicted to phonetically enhance their phonological features. Two other vowels also lowered, /ɛ/ and /ʌ/, which were phonologically mid vowels. Although no other vowels lowered, the high lax vowels /ɪ/ and /ʊ/ failed to raise, despite other phonologically high vowels /i/ and /u/ and tense mid vowels /e/ and /o/ doing so. One way to interpret this is in line with the prediction of Lee et al (2006), that sonority expansion more strongly affects lax vowels. This could alternatively be described phonetically as the effect of sonority expansion weakening for vowels located further from the bottom or center of the vowel space. The failure of phonologically high vowels /ɪ/ and /ʊ/ to rise may then reflect an equilibrium between the demands of sonority expansion and contrast enhancement. However, while /ɛ/ and /ʌ/ are phonologically mid vowels, they were quite low phonetically, both being well below the midpoint of the vowel space. Their lowering is thus also compatible with a more phonetically driven prominence strategy, by which vowels move away from the center of the vowel space rather than moving strictly to enhance their phonological features.

In contrast to the predicted effects of sonority expansion but in line with those of contrast enhancement, the effects of prominence were found to be bidirectional, bivalent, and sensitive to vowel identity. These effects were not, however, consistent across vowel categories in that not all vowels with shared phonological features patterned in the same way. In particular, the relationship between vowel height and movement along F1 was not straightforward. Mid vowels did not pattern together along this dimension: Tense mid vowels raised while lax mid vowels lowered, neither of which can be straightforwardly interpreted as enhancement of a "mid" vowel feature. Similarly, tense high vowels behaved differently than their lax counterparts. Lax vowels /ʊ/ and /ɪ/ only moved along F2, while their tense counterparts moved along either only F1, in the case of /i/, or along both dimensions, in the case of /u/.

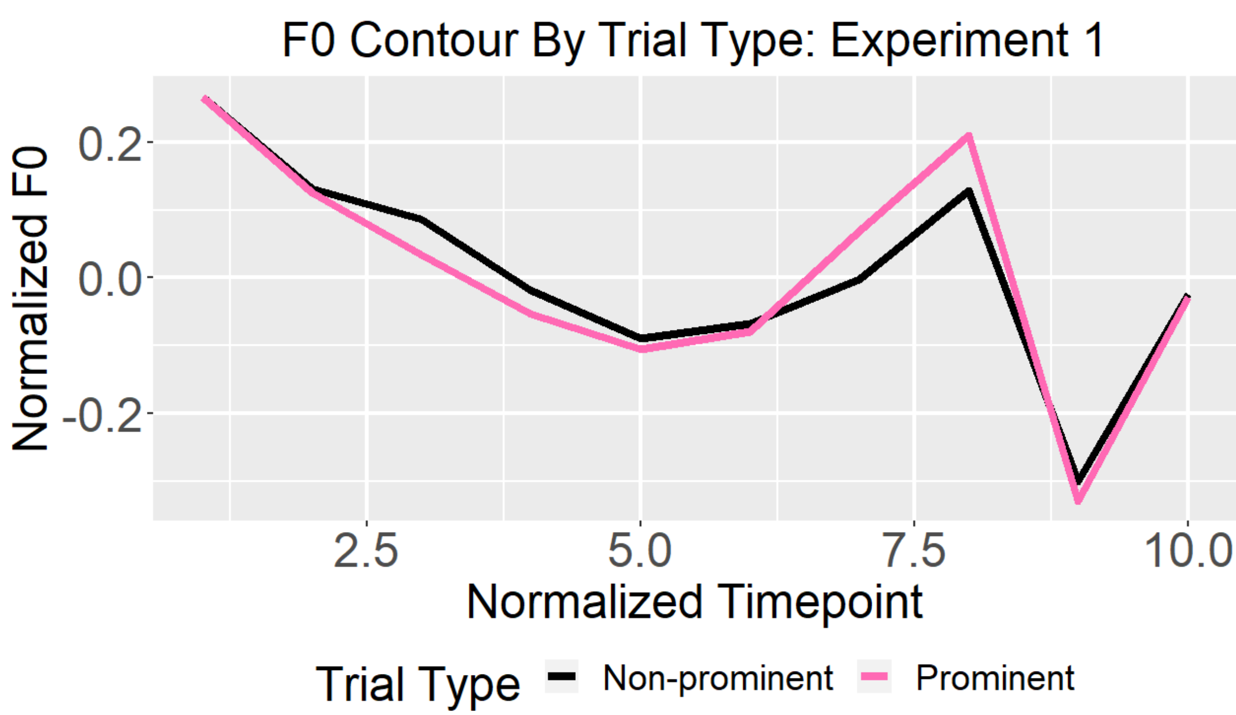
However, although not all vowels that belong to a shared phonological class patterned together, vowels that patterned together tended to share phonological features such that the effects of prominence were symmetrical across the vowel space along the front-back dimension. That is, both low vowels (/æ/ and /a/) lowered with no horizontal movement. Both lax mid vowels (/ɛ/ and /ʌ/) lowered and peripheralized horizontally; tense mid vowels (/e/ and /o/) raised and peripheralized horizontally. High lax vowels (/ɪ/ and /ʊ/) shifted horizontally with no movement along F1. The sole exception to this pattern along monophthongs was found in tense high vowels, where /u/ moved along F1 and F2 while /i/ only moved along F1.

The failure of all vowels to move in all directions was a prediction of the vowel extrinsic contrast enhancement hypothesis. However, contrary to this hypothesis, vowel movement did not appear to be constrained by the presence of a phonological neighbor along the relevant dimension. While the failure of /ɪ/ and /ʊ/ to raise would be, on its own, explainable by a vowel extrinsic hypothesis as avoiding the possibility of creating confusion with their tense counterparts /i/ and /u/, this would not explain why both vowels *did* peripheralize along F2. Similarly, both of the tense mid vowels /e/ and /o/ raised under focus, potentially bringing their midpoints closer to those of the corresponding high tense vowels. In summary, the finding that some vowels shift along only one dimension, but that those dimensions do not seem to correlate with their own phonological properties or those of their neighbors, is unexpected and does not straightforwardly accord with any hypothesized prominence strategies.

The behavior of the diphthongs /oɪ/, /aɪ/, and /aʊ/ was generally comparable to that of tense monophthongs and in line with peripheralization effects. Two of the vowels which failed to move along F1 at midpoint, /oɪ/ and /aʊ/ were found to raise and lower respectively at their 20% duration marks. This is in line with the behavior of the phonologically and phonetically similar /o/ and /a/ vowels. The behavior of /aɪ/ and /oɪ/ along F2 was inconsistent despite their shared status as fronting diphthongs. Only /oɪ/ moved further back along F2 at its 20% mark. It was also the sole vowel to move in opposite directions at its 20% and 80% marks. This means that /oɪ/ shifted in line with the phonological and phonetic characteristics of its constituent parts. That is, the back /o/ backed and front /ɪ/ fronted. However, this effect was not found for /aɪ/. This vowel is thus the sole diphthong that did not shift along F2 in either direction at any timepoint. This is not readily explained by articulatory constraints, as /aʊ/ occupies a similar position at midpoint but did exhibit backing under prominence. In short, diphthongs were similar to monophthongs in that phonological similarity did not guarantee that vowels would pattern together with regards to prominence.

Lack of an effect of focal prominence on F0 at midpoint also calls for an explanation. It is possible that the non-credible differences between these conditions indicates that participants were marking all critical words as prominent regardless of information status. This seems plausible, given that participant attention was always drawn towards the object of the sentence regardless of trial type. Figure 6 offers some qualitative support for the hypothesis that objects were always under focus. The y-axis in this Figure represents F0, which has first been scaled across speakers and then averaged across all speakers for both trial types. This F0 measurement extends across the entire sentence, with the x-axis showing normalized time as a proportion of the length of the utterance. F0 measurements were taken at 10% duration points for all sentences. Note that tokens for which F0 measurements could not be recovered at any such timepoint were excluded from this analysis, and thus fewer tokens are represented than were in the statistical analysis. As this Figure 5 illustrates, the mean pitch contours for each condition, while not identical, were quite similar in the height and location of their F0 maximum. It is therefore possible that the formant and duration differences observed in this study represent differences between different sources of phonetic prominence, such as between narrow and corrective focus. While the present experiment shows that focal prominence was more reliably encoded in duration and segmental peripheralization than F0, it provides no direct explanation as to why this would be the case. The null F0 results could also be explained if the effects attributed to focal prominence in fact are due to some other process favoring hyperarticulation, such as localized clear speech.

Figure 5: Time-normalized F0 contours for prominent and non-prominent productions, Experiment 1



Conclusion

The present experiment investigated the types of phonetic alterations vowels undergo under focal prominence. It expands on previous investigations of this phenomenon in its use of a relatively large scale, experimentally controlled method and in its inclusion of vowels from across the English vowel space. By analyzing the effects of prominence across the majority of American English vowels, it was possible to observe what effects of prominence appeared to be consistent across vowels and which varied between phonological categories. I argue that the observed differences between prominent and non-prominent productions are most clearly in line with the predictions of a contrast enhancement account. A more immediate goal of this experiment is to provide hypotheses for the next experiment presented in Experiment 2, which utilizes a very similar paradigm to ask how the effects of focal prominence extend to an artificial language context. The use of an artificial language is critical for the experiments that follow, as they require precise control over the distribution of phonetic information associated with a vowel category and over the social evaluation attached to a specific variant. Comparing the results of Experiment 1 to those of Experiment 2 will thus serve as a test of ecological validity of the remaining studies of this dissertation.

Chapter 3: Prominence in an Artificial Language

Experiment 2: Introduction

Experiment 2 is similar to Experiment 1 in that both are production experiments designed to elicit prominence by relying on its relationship to corrective focus (hereafter, simply referred to as *focus*). It differs from Experiment 1 in making use of an artificial language-learning paradigm meant to mimic the types of L2 learning exercise a learner might encounter in a classroom environment. Like Experiment 1, the experiment consists of a series of “prominent” which are marked by focus, and “non-prominent” trials which are not. Rather than correcting a simulated computer learner’s guesses at describing a simple situation, participants in Experiment 2 are asked to aid a virtual language learner by correcting or confirming their use of vocabulary in the artificial language. On critical trials, participants are exposed to a set of visual stimuli which corresponds to a target phrase, similarly to Experiment 1. These visuals are accompanied by audio of a “student,” of the language attempting to produce the corresponding phrase. On “non-prominent,” trials, the student produces the correct phrase and the participant confirms their guess by repeating it. On “prominent,” trials, the student substitutes one word in the target phrase with another word from the artificial language. In these contexts, participants correct their guess by producing the correct word, using prosodic prominence to mark corrective focus on that word.

Methods

Participants

A total of fifty-two participants were recruited remotely via Prolific. Data from twenty-two was excluded due to technical issues, poor audio quality, or failure to follow instructions. This resulted in a total of thirty participants whose data were analyzed. Participants’ mean age was 39, standard deviation of age was 9.36 and all reported growing up monolingual in the United States. One reported having spent three years studying Japanese. Regional representation was roughly comparable to Experiment 1. Nine participants were coded as being from the South (Florida, North Carolina, Texas, Alabama, West Virginia, Arkansas and Kentucky) nine from the Midwest (Michigan, Missouri, Kansas, Minnesota, Iowa and Ohio), seven from the Northeast (New York, New Jersey, Pennsylvania and Massachusetts) and four from the West (California and Washington). Fifteen identified as male, fourteen as female, and one as non-binary.

Stimuli

Auditory stimuli were recordings of nonce words of the form $C_1V_1C_2V_2$. Both consonants were always obstruents from the set {b, p, d, k, g, f, v, s, z}. V_1 was always from the set of critical vowels, which were divided into a set of "tense" vowels {i, e, u, a} and a set of relatively "lax" vowels {ɪ, ɛ, o, ʌ}. /o/ was included with the lax vowels rather than /ʊ/ despite being a phonologically tense vowel. This was done because of the relatively low type frequency of /ʊ/ in English. In addition, I hypothesized that the dearth of English minimal pairs differentiated by /ʊ/ and /u/ could potentially influence results of Experiments 3 and 4 by biasing speakers towards more /u/ like productions, so /ʊ/ was removed to maintain comparability between these three experiments. V_2 was always schwa. Stimuli consisted of sixteen lexical items. This is divided into two sets, set A and set B. Words in each set were identical except in the critical vowel. Set A consisted of eight words containing critical peripheralized vowels while words in set B contained the corresponding centralized vowels. Each set was further divided into two non-overlapping subsets: A_1 (/guvə/, /pibə/, /sepə/, /zabə/), B_1(/vopə/, /dʌzə/, /kɛbə/, /zɪfə/) A_2 (/vupə/, /dazə/, /kebə/, /zifə/), and B_2(/gova/, /pɪbə/, /sɛpə/, /zʌbə/) Each critical vowel appeared in one word per subset

Model talker recordings consisted of two sets of training recordings and two sets of test recordings. Training productions were recorded by a model talker identified in the experiment as the "teacher." One set of training productions consisted of all words in the language spoken in isolation. The other training set consisted of recordings of short phrases of the form "Word_1 to Word_2'" where both words were from the same subset. There were eight such recordings. Each word in a set appeared twice as Word_1 and twice as Word_2. Identically structured recordings were made by the two other speakers.

Each speaker recorded every possible phrase that could contain two distinct words from the same set. This resulted in a total of twelve recordings across four sets per speaker for a total of ninety-six recordings in all. Model talkers were all female L1 speakers of American English in their mid-twenties to thirties. One model talker identified as Asian American and as having grown up in California. One identified as mixed Asian and white, also having grown up in California. Two reported growing up primarily in the Midwest, in Ohio and Iowa, and identified as white.

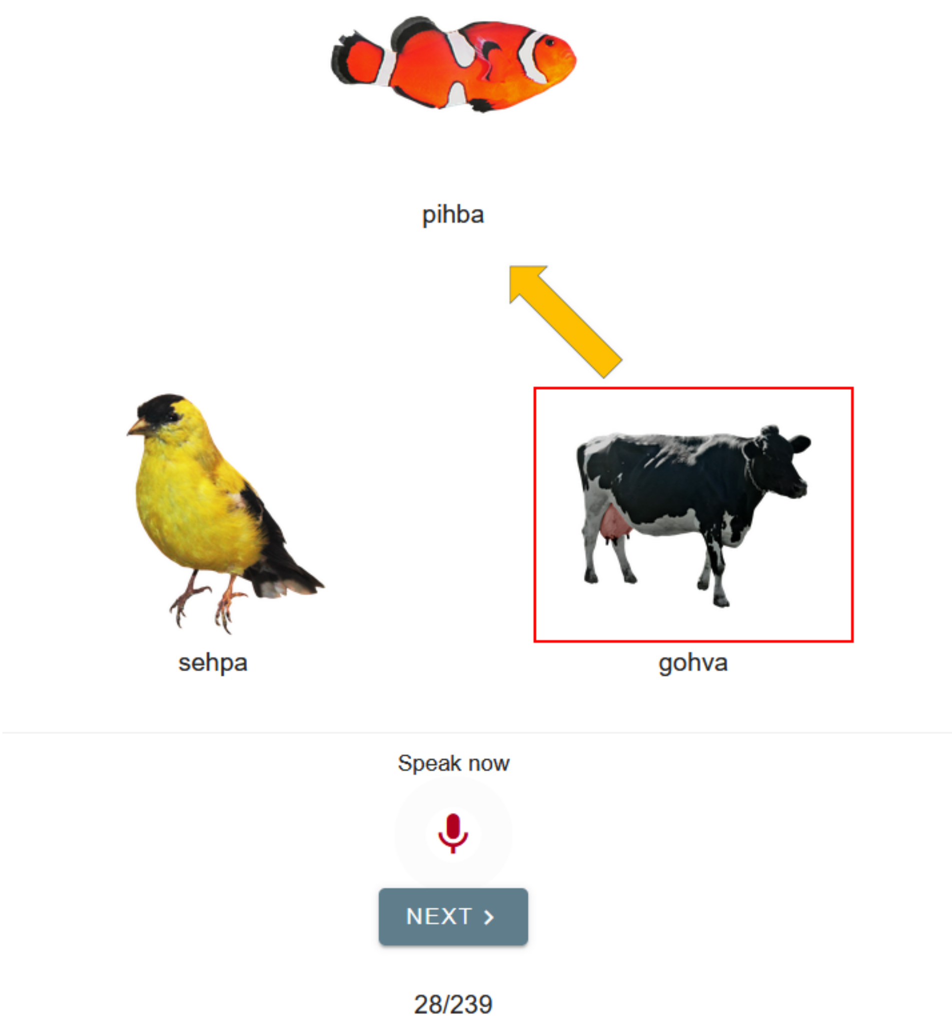
Visual stimuli consisted of a set of eight images from the Bank of Standardized Stimuli (Brodeur et al., 2010). These included four pictures of animals (a canary, a dalmatian, a clownfish and a cow) and four of food items (a cookie, a stock of broccoli, a banana and a lemon). Orthographic representations of each word in the artificial language appeared under each picture on each trial. The vowels /i/, /ɪ/, /e/, /ɛ/, /o/, /u/, /ʌ/ and /a/ were represented by "ee," "ih", "ay", "eh", "oh", "oo", "ah" and "aa" respectively. Final schwa was represented by word final "a." Each image was consistently assigned to either one word from set A or the corresponding word from set B depending on condition.

Procedure

As was the case for Experiment 1, Experiment 2 was conducted remotely through the use of a web application. The experiment consisted of two identically structured blocks. Each block contained two training sections and a test section. Participants were randomly assigned to one of three conditions. In condition one, participants were exposed to tense set A_1 in the first block and lax B_1 in the second. Participants in condition two heard A_2 followed by B_2. Participants in condition three heard B_1 followed by A_1. Given the limited number of participants available, these three conditions allowed both block orders and both vocabulary sets to be represented. Ten participants were assigned to each condition.

All participants were first given an overview of the experiment. They were told that they would be learning words in a new language and shown a spelling guide. They then proceeded to the first training round in which they were instructed to repeat after a model "teacher." Participants were then exposed to one subset of four words. The teacher produced each of these four words twice, each time accompanied by the appropriate image. The order of examples was randomized between participants. Participants moved on to the next trial by clicking a button at the bottom of the screen. There was no time limit for each trial. Following this, they moved on to the second training round in which the same teacher produced short phrases of the form "$Word_1$ to $Word_2$" These phrases were accompanied by visual stimuli consisting of three images arranged in a triangle. The two images at the bottom of the triangle represented potential subjects of the target phrase ($Word_1$) while the image on top represented the object of the phrase ($Word_2$). At the beginning of each trial, a yellow arrow appeared pointing from one of the potential subject images to the object image. This represented the target phrase in which $Word_1$ was always the image at which the arrow started. Participants completed eight such trials and the order of trials was randomized between participants.

Figure 6: Illustration of a "Subject Prominent" trial



Following the training phase, participants were given a new set of instructions introducing two new model talkers, identified as “fellow students.” They were told that their role would be to help these students study the language. On each round, one of the students would attempt to produce the target phrase represented by the images. The student would either produce the correct phrase (non-prominent trial) or a phrase in which either Word_1 (subject prominent trial) or Word_2 (object prominent trial) was incorrect. The errorful word was always the word from the set of four critical words in the same block which was not visually presented on that trial. Unlike Experiment 1, participants could be asked to correct either the subject or the object of the phrase as opposed to only the object. Figure 6 illustrates an example of a subject prominent trial. In this case, the spoken phrase would be “zaaba to pihba” and the target phrase would be “gohva to pihba.” When the student produced the correct phrase, the participant was simply instructed to repeat it back to them. When the student produced an incorrect phrase, the participant was told to correct the student by saying the correct phrase and communicating that the answer was incorrect “using their tone of voice.” In addition, the errorful word on prominent trials was marked by a red outline.

Figure 7: Block and round structure of Experiment 2. Note that the order of trials in test rounds was randomized between participants on a per-trial basis.

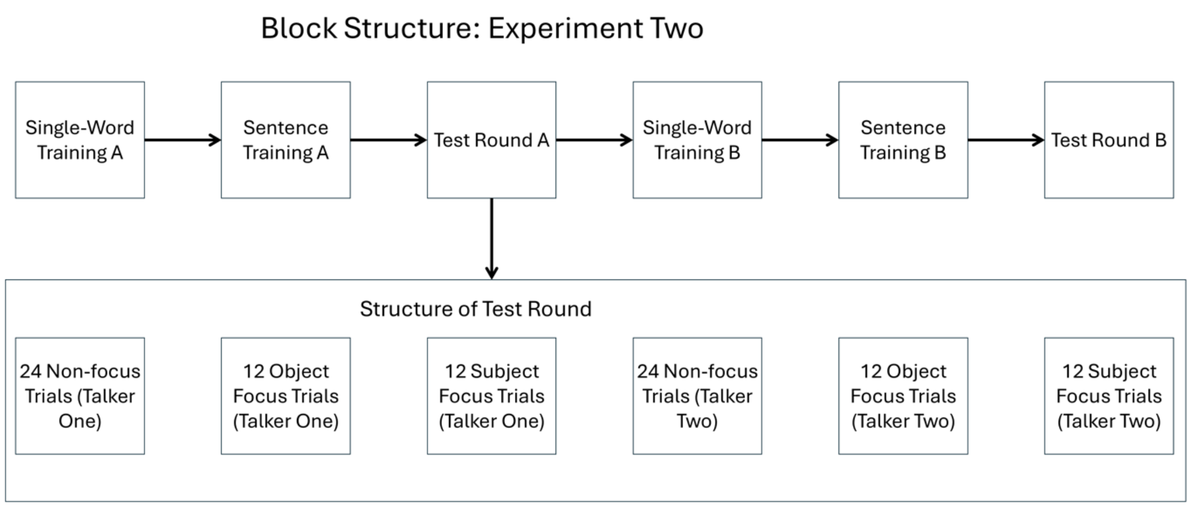


Figure 7 illustrates the overall structure of the experiment. There were ninety-six test trials per block: forty-eight non-prominent, twenty-four object prominent, and twenty-four subject-prominent. After completing ninety-six trials, participants were asked to take a brief break before moving onto the next block. This block was identical to the first except for the vocabulary items, which were changed as outlined above (participants who heard set A in the first round were exposed to set B in the second and vice versa). There were twelve possible sentences in each block: One for each combination of unique words. For example, Word A could appear as the subject in “Word A to Word B,” “Word A to Word C” and “Word A to Word D.” Each possible sentence appeared as the target sentence four times in non-prominent conditions (twice per model talker), twice in subject-prominent conditions (once per model talker) and twice in object-prominent conditions. [i]

Analysis

Analyses were similar to Experiment 1. Each trial yielded one subject and one object word. Midpoint F1 and F2 values were statistically analyzed using a multivariate Bayesian mixed-effects regression model (N = 11,010). Fixed effects were defined for vowel (an eight-way predictor), sentence position (subject or object) and prominenc condition as well as their interactions. Prominence condition was encoded as a three-way predictor with levels non-prominent (no word in the sentence was under corrective focus), prominent (the word analyzed was under corrective focus) and deaccented (the other word in the sentence was under corrective focus). Random intercepts were defined by participant and random slopes defined for all fixed effects. Log duration and midpoint F0 were analyzed using identical models. As was the case in Experiment 1, trials for which F0 could not be reliably measured were excluded from the F0 analysis (N = 10,749). Given that /e/ and /o/ were the only diphthongs analyzed and given that these behaved similarly to monophthongs in the previous experiment, no additional models were run analyzing different timepoints of these categories. In addition to reporting 95% CI as in the last experiment, several effects were credible and reported at the 89% CI level. These are another common credible interval for Bayesian analysis and can be thought of as roughly analogous to marginally significant effects in frequentist models. As in Experiment 1, credible effects were defined as those whose credible interval did not cross zero.

In addition to the effects of prominence analyzed in this experiment and Experiment 2, the fact that productions included subject-prominent and object-prominent conditions meant that it was possible to measure the effect of “deaccenting,” as well. This refers to the differences between the production of vowels when they occur in confirm sentences and when they occur in sentences in which the opposite constituent is focused: Subjects in object-prominent sentences and objects in subject-prominent sentences. The term “deaccented” therefore describes pre-focal material in object-prominent trials and post-focal in subject-prominent trials.

Results

Figure 8: Average effect of prominence across subject and object positions. Pink letters represent means for prominent trials, black letters represent means for confirm trials

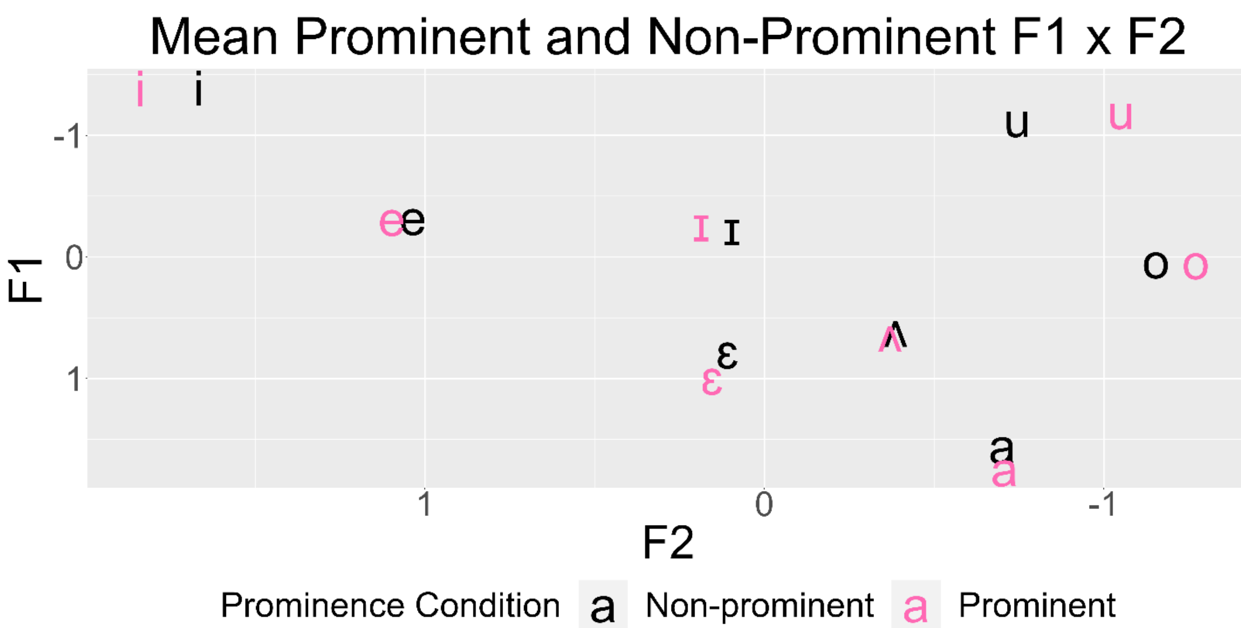
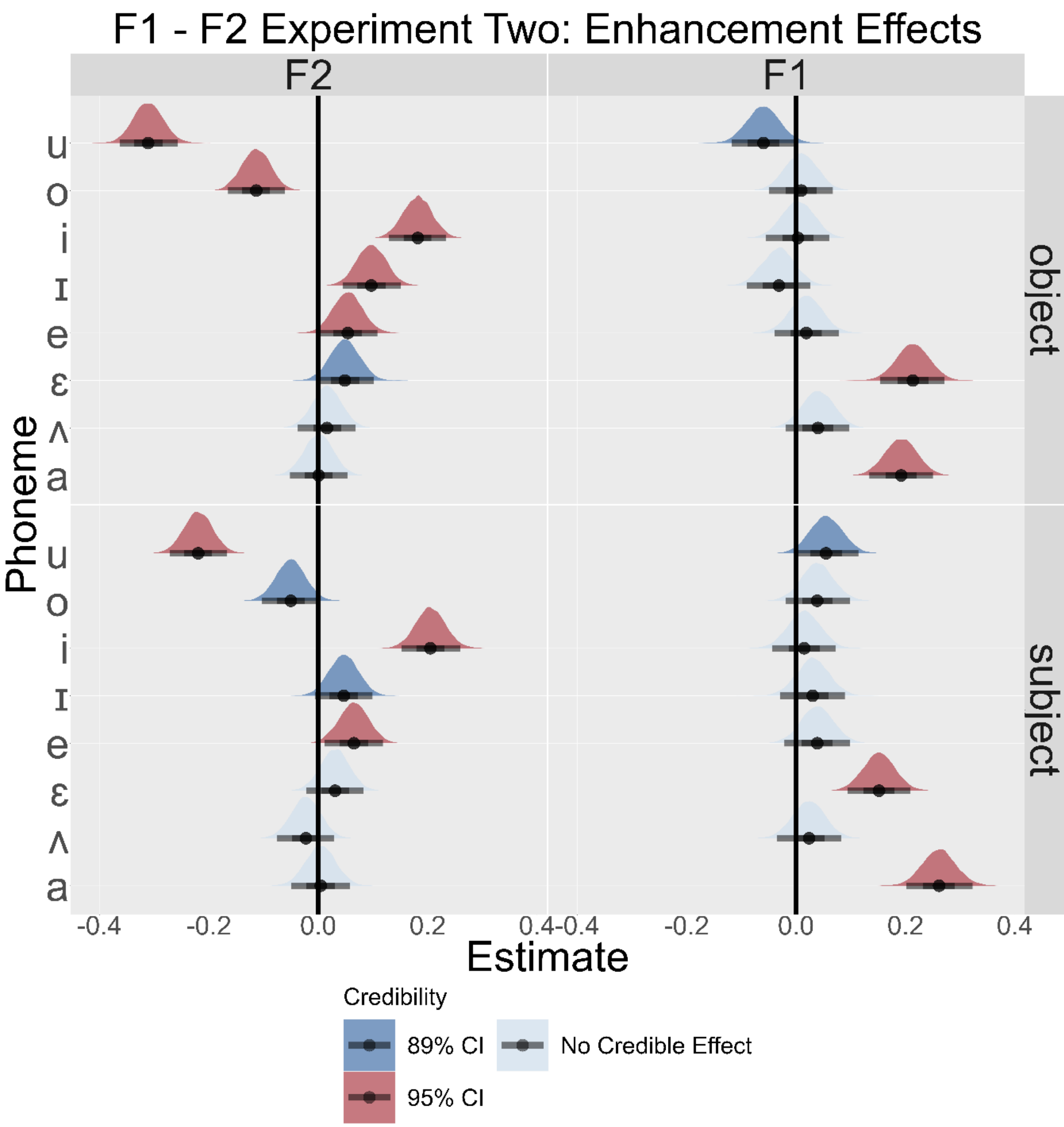


Figure 8 illustrates the mean normalized F1 and F2 values of all vowel categories, averaged across sentence positions. The effects of enhancement for each vowel category in each sentence position are illustrated in Figure 9. Each distribution represents the modeled value of the effect in question. In the case of F2, values further to the right indicate greater fronting under prominence and values to the left greater backing. For F1, values to the left represent phonetic raising under prominence and those to the right represent phonetic lowering. Zero is represented by the black vertical line and effects are considered credible if their associated confidence interval does not cross this line at either the 89% or 95% level. Red distributions represent effects that are credible at the 95% level, dark blue those at the 89% level, and light blue non-credible effects. As an example, the distribution representing the effect of prominence on F2 of /ɪ/ is red in the object position, indicating that this vowel credibly fronts under prominence in this condition at the 95% credible level. The corresponding F2 distribution in subject position is blue, which indicates that the effect is credible only at the 89% level. The distributions representing the effect of prominence on F1 of /ɪ/ are light blue in both subject and object position, indicating that the effect of prominence did not reach credibility at either level.

Figure 9: Effect of prominence by sentence position and vowel



A general trend emerges where vowel height is related to the magnitude of the effect of prominence on F2. In object position at the 95% level, /u/, /i/, /ɪ/, /o/ and /e/ all show credible movement along this dimension. The movement of each vowel is in line with what would be predicted by their backness specification. Looking first at effects in the object position (top panel of Figure 9), the effect of prominence on F2 of /u/ is negative, meaning that this vowels phonetically backed under corrective focus. The F2 coefficient of /i/ is positive, though smaller in magnitude than that of /u/, indicating that the vowel fronted under corrective focus to a lesser degree than /u/ backed. This horizontal movement decreased in magnitude as vowels phonetically and phonologically lowered. Mid back /o/ showed the next largest effect compared to /i/ and /u/, followed by /ɪ/ and /e/ respectively. A credible fronting effect was also observed for /ɛ/ in this position, though only at the 89% CI. A similar trend can be observed in subject position, though the relationship between vowel height and prominence effects on F2 is weaker and less consistent. Again here, /u/ and /i/ show the largest effects of prominence. A credible fronting effect of prominence is also seen for /e/. Unlike object position, the effect of prominence on F2 of /o/ and /ɪ/ is credible only at the 89% level and no credible effect was observed for /ɛ/. Movement along F1 was also more limited. Robust lowering effects were observed for /ɛ/ and /a/ in both subject and object positions. The only other vowel to undergo movement along F1 in subject position was /u/, which was also the only vowel to move in contradictory directions in subject and object positions. /u/ credibly raised under prominence in object position but lowered in object position. Both effects on /u/ were visible only at the 89% level. The vowel /ʌ/ did not credibly move along either dimension.

Table 6: Posterior Estimates of Enhancement Effects on Vowels: F1

	Object		Subject	
Phoneme	**β**	**CI**	**β**	**CI**
a	0.192	**0.95 (0.131, 0.247)**	0.261	**0.95 (0.202, 0.323)**
		0.89 (0.144, 0.239)		**0.89 (0.213, 0.31)**
ʌ	0.040	0.95 (-0.017, 0.099)	0.024	**0.95 (-0.034, 0.082)**
		0.89 (-0.005, 0.089)		**0.89 (-0.025, 0.069)**
ɛ	0.214	**0.95 (0.154, 0.271)**	0.152	**0.95 (0.094, 0.209)**
		0.89 (0.168, 0.264)		**0.89 (0.105, 0.199)**
e	0.019	0.95 (-0.041, 0.076)	0.039	0.95 (-0.019, 0.1)
		0.89 (-0.03, 0.066)		0.89 (-0.008, 0.088)
ɪ	-0.031	0.95 (-0.089, 0.027)	0.031	0.95 (-0.029, 0.089)
		0.89 (-0.079, 0.016)		0.89 (-0.019, 0.077)
i	0.004	0.95 (-0.057, 0.06)	0.015	0.95 (-0.044, 0.072)
		0.89 (-0.044, 0.051)		0.89 (-0.035, 0.06)
o	0.010	0.95 (-0.048, 0.068)	0.039	0.95 (-0.022, 0.096)
		0.89 (-0.038, 0.056)		0.89 (-0.008, 0.087)
u	-0.060	0.95 (-0.118, 0)	0.055	0.95 (-0.002, 0.114)
		0.89 (-0.107, -0.011)		**0.89 (0.008, 0.103)**

Table 7: Posterior Estimates of Enhancement Effects on Vowels: F2

	Object		Subject	
Phoneme	**β**	**CI**	**β**	**CI**
a	0.001	0.95 (-0.053, 0.051)	0.005	0.95 (-0.047, 0.061)
		0.89 (-0.042, 0.043)		0.89 (-0.04, 0.048)
ʌ	0.016	0.95 (-0.037, 0.069)	-0.023	0.95 (-0.074, 0.03)
		0.89 (-0.025, 0.06)		0.89 (-0.065, 0.02)
ɛ	0.049	0.95 (-0.004, 0.101)	0.031	0.95 (-0.021, 0.083)
		0.89 (0.006, 0.092)		0.89 (-0.011, 0.073)
e	0.054	**0.95 (0, 0.107)**	0.065	**0.95 (0.011, 0.117)**
		0.89 (0.01, 0.097)		**0.89 (0.021, 0.109)**
ɪ	0.097	**0.95 (0.044, 0.15)**	0.047	**0.95 (-0.005, 0.099)**
		0.89 (0.055, 0.14)		**0.89 (0.005, 0.09)**
i	0.182	**0.95 (0.129, 0.232)**	0.205	**0.95 (0.15, 0.257)**
		0.89 (0.141, 0.226)		**0.89 (0.16, 0.248)**
o	-0.113	**0.95 (-0.166, -0.064)**	-0.050	0.95 (-0.102, 0.004)
		0.89 (-0.156, -0.072)		**0.89 (-0.093, -0.007)**
u	-0.311	**0.95 (-0.363, -0.257)**	-0.219	**0.95 (-0.271, -0.166)**
		0.89 (-0.354, -0.268)		**0.89 (-0.263, -0.178)**

The effect of deaccenting was inconsistent between vowel categories and sentential positions. Contrary to expectations, most vowels that moved under deaccenting did so in the same direction in which they shifted under prominence. Only /ʌ/ moved differently in the two conditions, raising in the deaccenting condition and not moving at all under prominence. All other vowels moved in the same direction under deaccenting as under prominence. Furthermore, the effect was inconsistent between sentence positions with no vowel showing credible deaccenting effects in both subject and object position. Only /o/ moved in the object position along F2 at the 95% level while /i/ and /e/ fronted in the subject position at the 89% level. In addition to the raising movement observed for /ʌ/, both /ɛ/ and /a/ lowered under deaccenting in object and subject, respectively.

Figure 10: Effect of Deaccenting across formants and sentence positions

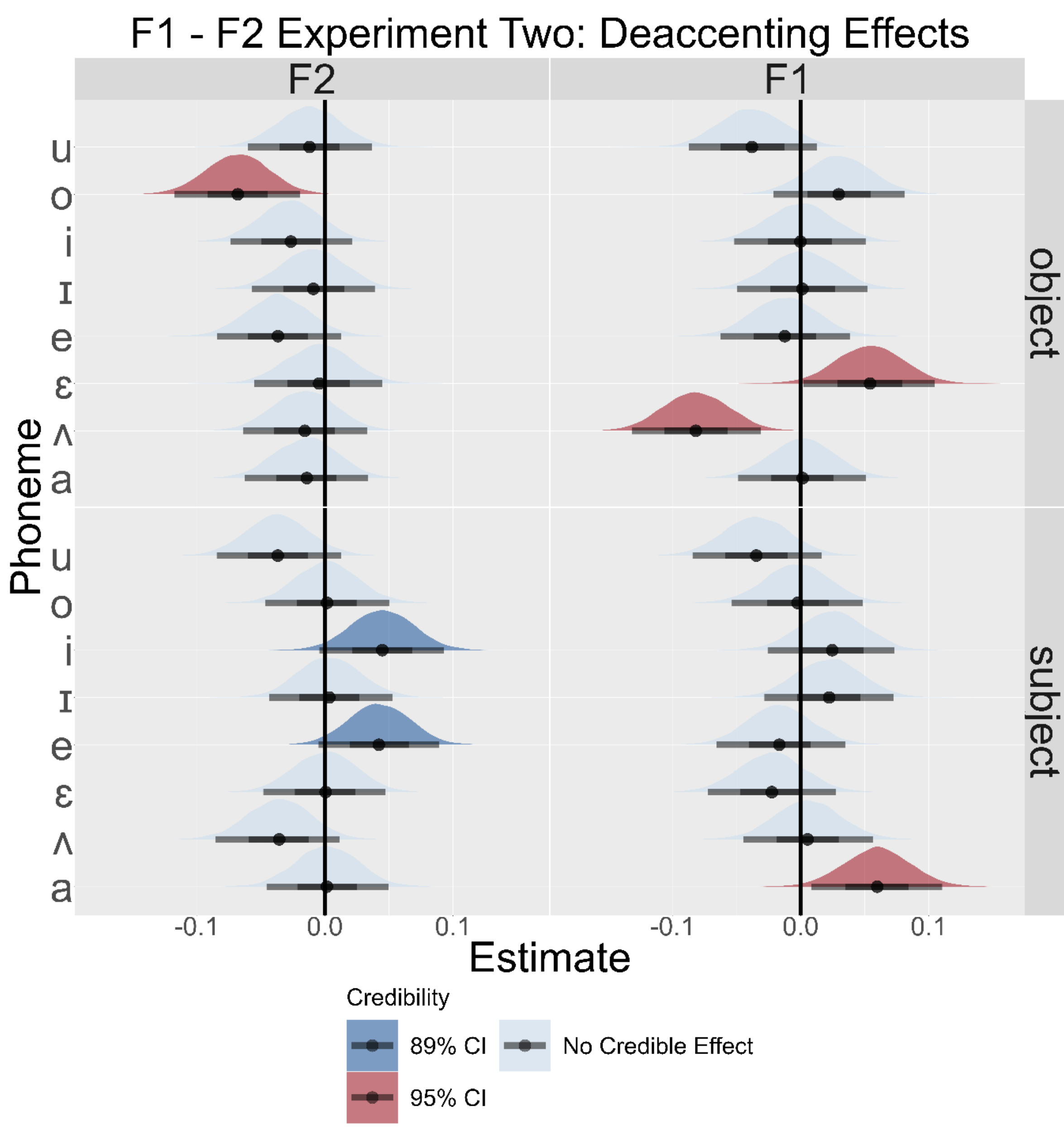


Table 8: Posterior Estimates of Deaccent Effects on Vowels F1

	Object		Subject	
Phoneme	β	CI	β	CI
a	0.002	0.95 (-0.05, 0.05)	0.060	**0.95 (0.009, 0.111)**
		0.89 (-0.04, 0.04)		**0.89 (0.019, 0.102)**
ʌ	-0.082	**0.95 (-0.133, -0.033)**	0.005	0.95 (-0.045, 0.055)
		0.89 (-0.121, -0.04)		0.89 (-0.036, 0.046)
ɛ	0.054	**0.95 (0.002, 0.105)**	-0.022	0.95 (-0.072, 0.028)
		0.89 (0.013, 0.096)		0.89 (-0.062, 0.02)
e	-0.012	0.95 (-0.06, 0.04)	-0.017	0.95 (-0.067, 0.034)
		0.89 (-0.051, 0.03)		0.89 (-0.055, 0.025)
ɪ	0.001	0.95 (-0.05, 0.052)	0.022	0.95 (-0.028, 0.073)
		0.89 (-0.04, 0.043)		0.89 (-0.021, 0.061)
i	0.000	0.95 (-0.053, 0.048)	0.025	0.95 (-0.025, 0.073)
		0.89 (-0.044, 0.04)		0.89 (-0.016, 0.066)
o	0.030	0.95 (-0.021, 0.082)	-0.003	0.95 (-0.056, 0.046)
		0.89 (-0.011, 0.072)		0.89 (-0.043, 0.039)
u	-0.038	0.95 (-0.088, 0.012)	-0.035	0.95 (-0.083, 0.017)
		0.89 (-0.079, 0.004)		0.89 (-0.076, 0.006)

Table 9: Posterior Estimates of Deaccenting Effects on Vowels: F2

	Object		Subject	
Phoneme	**β**	**CI**	**β**	**CI**
a	-0.014	0.95 (-0.063, 0.033)	0.002	0.95 (-0.046, 0.049)
		0.89 (-0.054, 0.023)		0.89 (-0.039, 0.039)
ʌ	-0.016	0.95 (-0.066, 0.03)	-0.036	0.95 (-0.084, 0.013)
		0.89 (-0.055, 0.023)		0.89 (-0.074, 0.004)
ɛ	-0.004	0.95 (-0.055, 0.044)	0.001	0.95 (-0.047, 0.048)
		0.89 (-0.046, 0.035)		0.89 (-0.039, 0.038)
e	-0.037	0.95 (-0.085, 0.012)	0.042	0.95 (-0.004, 0.09)
		0.89 (-0.078, 0.001)		**0.89 (0.005, 0.081)**
ɪ	-0.009	0.95 (-0.057, 0.038)	0.003	0.95 (-0.044, 0.052)
		0.89 (-0.047, 0.032)		0.89 (-0.033, 0.045)
i	-0.027	0.95 (-0.074, 0.022)	0.045	0.95 (-0.004, 0.093)
		0.89 (-0.067, 0.011)		0.89 (0.006, 0.084)
o	-0.068	**0.95 (-0.117, -0.02)**	0.002	0.95 (-0.045, 0.051)
		0.89 (-0.109, -0.03)		0.89 (-0.038, 0.039)
u	-0.012	0.95 (-0.06, 0.037)	-0.037	0.95 (-0.083, 0.014)
		0.89 (-0.051, 0.028)		**0.89 (-0.077, 0.002)**

Both F0 and duration showed credible effects of prominence-related enhancement and reduction when aggregated across vowels. For both subject and object position, vowels were found to exhibit higher midpoint F0 under prominence compared to confirm conditions (objects: $\beta = 0.564$, 95% CI = [0.395, 0.729]; subjects: $\beta = 0.569$, 95% CI = [0.401, 0.731]). Contrary to expectations, F0 in subject position was also higher under deaccenting compared to confirm conditions

($\beta = 0.118$, 95% CI = [0.01, 0.225]). This was not the case for objects, which showed credibly lower F0 under deaccenting ($\beta = -0.327$, 95% CI = [-0.436, -0.223]). For F0, these effects held to a similar degree across all vowel categories in both positions. The effect of duration was more variable across categories. The majority of vowels were longer under prominence compared to non-prominent conditions, in both subject and object positions. These effects were greater for subject position on average. However, /ɛ/, /ɪ/ and /o/ did not undergo credible lengthening under prominent in the object position. Comparing deaccented vowels with prominent and confirm conditions, only two vowels, /o/ and /a/, showed an effect of deaccenting, with shorter duration at the 89% and 95% levels. respectively.

Figure 11: Effect of Prominence on Vowel Duration. Model Estimates Represent Differences Between "Confirm" and "Prominent" Trials.

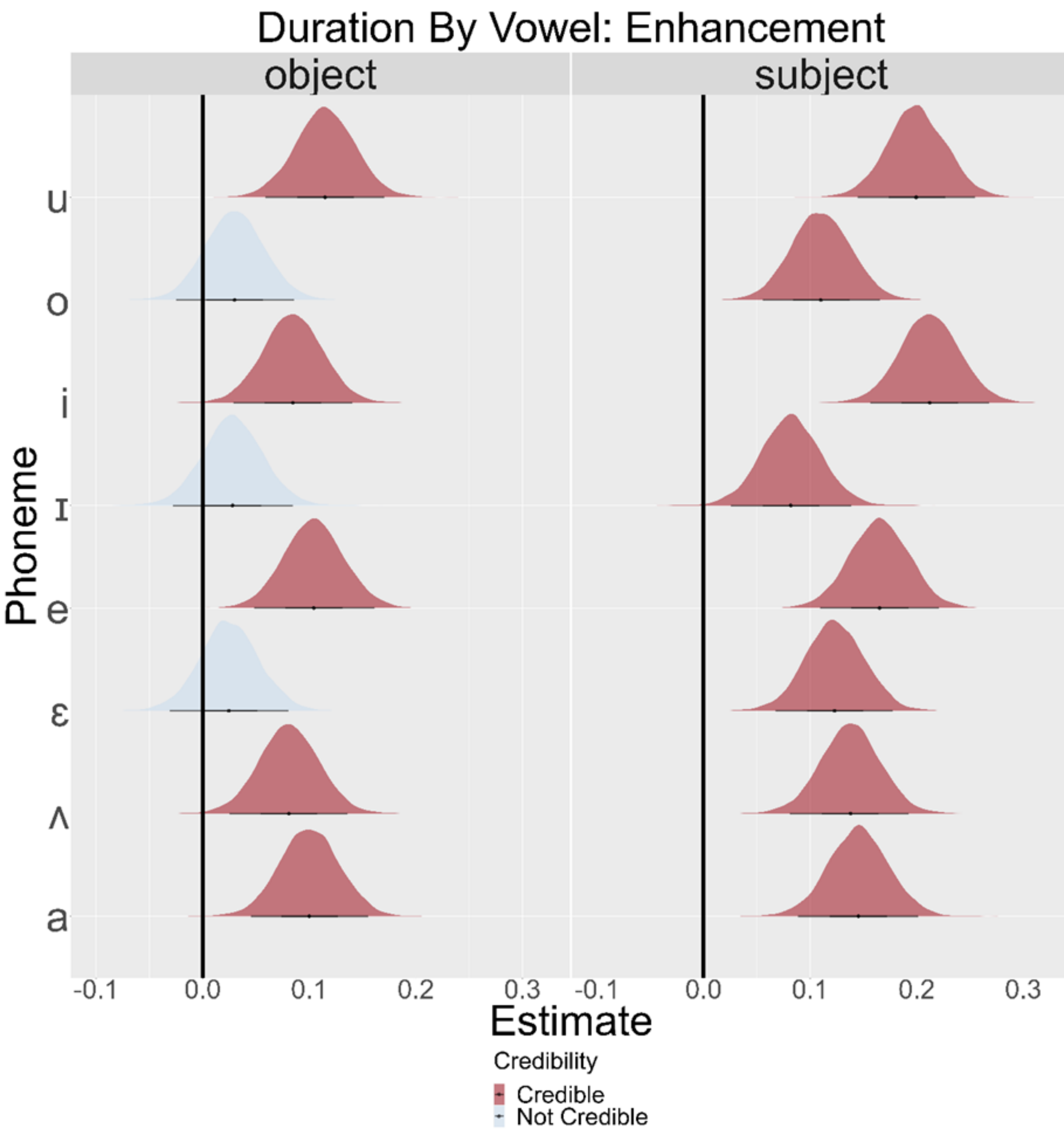


Figure 12: Effect of deaccenting on vowel duration. Model estimates represent differences between "confirm" and "deaccented" trials

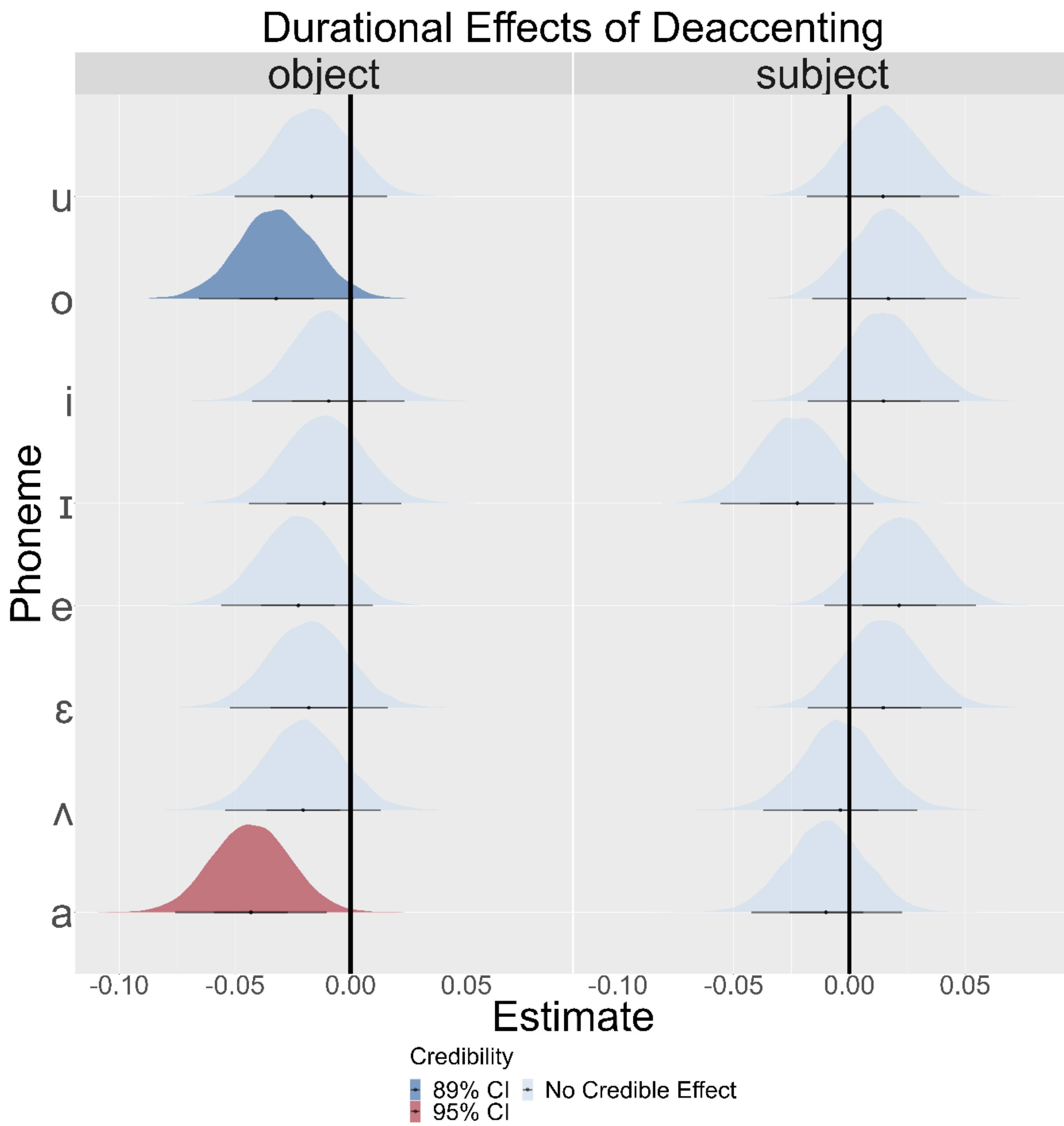


Table 10: Posterior Estimates of Enhancement Effects on Vowels: Log Duration

	Object		Subject	
Phoneme	**β**	**CI**	**β**	**CI**
a	0.100	**0.95 (0.047, 0.156)**	0.145	**0.95 (0.088, 0.2)**
		0.89 (0.055, 0.145)		**0.89 (0.102, 0.192)**
ʌ	0.081	**0.95 (0.025, 0.136)**	0.138	**0.95 (0.081, 0.192)**
		0.89 (0.035, 0.124)		**0.89 (0.092, 0.181)**
ɛ	0.025	0.95 (-0.03, 0.082)	0.123	**0.95 (0.068, 0.178)**
		0.89 (-0.018, 0.073)		**0.89 (0.081, 0.17)**
e	0.104	**0.95 (0.049, 0.161)**	0.165	**0.95 (0.11, 0.221)**
		0.89 (0.06, 0.15)		**0.89 (0.121, 0.211)**
ɪ	0.028	0.95 (-0.026, 0.086)	0.082	**0.95 (0.026, 0.139)**
		0.89 (-0.016, 0.074)		**0.89 (0.037, 0.127)**
i	0.085	**0.95 (0.029, 0.14)**	0.212	**0.95 (0.157, 0.268)**
		0.89 (0.039, 0.129)		**0.89 (0.168, 0.258)**
o	0.030	0.95 (-0.025, 0.086)	0.110	**0.95 (0.056, 0.166)**
		0.89 (-0.015, 0.074)		**0.89 (0.064, 0.153)**
u	0.115	**0.95 (0.057, 0.168)**	0.199	**0.95 (0.145, 0.255)**
		0.89 (0.071, 0.161)		**0.89 (0.157, 0.246)**

Table 11: Posterior Estimates of Deaccenting Effects on Vowels: Log Duration

	Object		Subject	
Phoneme	**β**	**CI**	**β**	**CI**
a	-0.043	**0.95 (-0.076, -0.011)**	-0.010	0.95 (-0.042, 0.023)
		0.89 (-0.07, -0.017)		0.89 (-0.037, 0.016)
ʌ	-0.021	0.95 (-0.055, 0.012)	-0.004	0.95 (-0.038, 0.029)
		0.89 (-0.048, 0.007)		0.89 (-0.031, 0.024)
ɛ	-0.018	0.95 (-0.051, 0.017)	0.015	0.95 (-0.019, 0.047)
		0.89 (-0.046, 0.009)		0.89 (-0.012, 0.042)
e	-0.023	0.95 (-0.056, 0.01)	0.022	0.95 (-0.011, 0.054)
		0.89 (-0.048, 0.006)		0.89 (-0.006, 0.047)
ɪ	-0.011	0.95 (-0.045, 0.021)	-0.022	0.95 (-0.056, 0.01)
		0.89 (-0.038, 0.016)		0.89 (-0.048, 0.005)
i	-0.009	0.95 (-0.042, 0.024)	0.015	0.95 (-0.018, 0.047)
		0.89 (-0.036, 0.018)		0.89 (-0.012, 0.042)
o	-0.032	0.95 (-0.066, 0.001)	0.017	0.95 (-0.016, 0.05)
		0.89 (-0.059, -0.005)		0.89 (-0.011, 0.043)
u	-0.017	0.95 (-0.05, 0.016)	0.015	0.95 (-0.017, 0.049)
		0.89 (-0.043, 0.011)		0.89 (-0.012, 0.041)

Discussion

As was the case for Experiment 1, the results of Experiment 2 suggest that phonetic correlates of contrastive focus are shaped by the phonetic and phonological characteristics of the prominent vowel in question. Like Experiment 1, vowels were generally found to peripheralize along F2 in the direction predicted by their phonological backness specification. Unlikely Experiment 1, however, prominence effects were mostly limited to F2. Results of Experiment 2 also suggest a roughly inverse relationship between vowel height and movement along F1. Only the two phonetically lowest vowel categories, /a/ and /ɛ/ exhibited consistent movement along this dimension across sentential positions. While /u/ was also found to move along F1, the direction of movement under prominence was inconsistent between sentential contexts and the magnitude of the effect much smaller than that of the lower vowels. Also unlike Experiment 1, in which all vowels shifted along at least one dimension, /ʌ/ did not credibly shift along either dimension under corrective focus. Overall, the results of Experiment 2 suggest speakers employ different strategies for marking prominence in the artificial language context compared to the natural language context.

On their own, the results of Experiment 2 provide stronger evidence of predominating effects of contrast enhancement and contrast maintenance compared to Experiment 1. Like Experiment 1, vowels phonetically peripheralized. However, rather than peripheralizing along all available dimensions, vowels shifted in ways that seem to maximize within-category distinctiveness while minimizing between-category confusability. Contrast maintenance, or a pressure to prevent vowels from shifting towards confusable neighbors, would provide an explanation for the observation that higher vowels tended to shift only along F2. Under this interpretation, mid and lax vowels may have failed to shift vertically because doing so would have been counterproductive to the goal of phonetic contrast enhancement. Specifically, raising the front vowels /e/ and /ɪ/ would bring both closer to /i/, increasing the opportunity for perceptual confusion. Similarly, raising/o/ would risk confusion with /u/, while lowering /ʌ/ would risk confusion with /a/. On the other hand, raising /u/ would not threaten contrast with any other vowel. As was the case for Experiment 2, the fact that /i/ and /a/ moved along only one dimension could potentially be attributed to articulatory constraints as both may have already been close to the periphery of the vowel space. Note however, that /i/ moved along F2 in Experiment 2 and failed to move along F1 while the reverse was true for Experiment 1. Finally, like Experiment 1, no evidence was found of vowels lowering in ways that did correlate with phonetic peripheralization, i.e., in ways that did not shift them away from the center of the vowel space. There is thus no evidence for speakers pursuing sonority expansion as a strategy, as this would predict that speakers lower all vowels regardless of phonetic or phonological specification.

The results from Experiment 2 provide evidence for a connection between sentential position and prominence effects on vowel formants. Specifically, prominence-related movement along F2 was less consistent between vowel categories in subject position compared to object. Vowels which did move along F2 in subject position also moved to a lesser degree than in object position. There is significantly less evidence for a relationship between sentence position and prominence-related movement along F1. The only vowel to behave differently in this position was /u/, which moved in opposite directions along F1 in phrase-initial and phrase-final contexts. In both cases, the magnitude of the observed effect was much smaller than that of the F1 effects showing lowering for /ɛ/ and /ʌ/. It has been hypothesized that sonority expansion serves to mark prosodic boundaries rather than prominence and this lowering may therefore function as a marker of an upcoming prosodic boundary (Cho, 2016). This explanation, however, seems unlikely given that the effect was restricted to /u/.

Conclusion

The goal of Experiment 2 was to establish how the phonetic correlates of prominence observed in the production of existing English words relate to the effects of prominence on novel words in an artificial language. Establishing the nature of prominence-related effects on vowel formants in the context of an artificial language is necessary because the remainder of the experiments described in this dissertation utilize an artificial language-learning paradigm. The most striking result of this experiment is that the effects of prominence in the artificial language context differed substantially from those observed in the natural language context. The correlates of prominence in the artificial language context were more restricted in that, with the exception of the vowel /u/, effects were observed either in F1 or

F2, but not generally in both dimensions. Still, prominence was consistently associated with phonetic peripheralization across both experiments. This suggests that the phonetic profile of prominence is at least in part predictable. This is important for establishing the ecological validity of the following experiments, as it supports the premise that the phonetic marking of pragmatic information, in this case focus, is comparable in the production of real words and words in an artificial language. This allows comparison with the phonetic marking of social information in the artificial language context.

Chapter 4: Study 1 Overall Discussion

Prominence Strategies

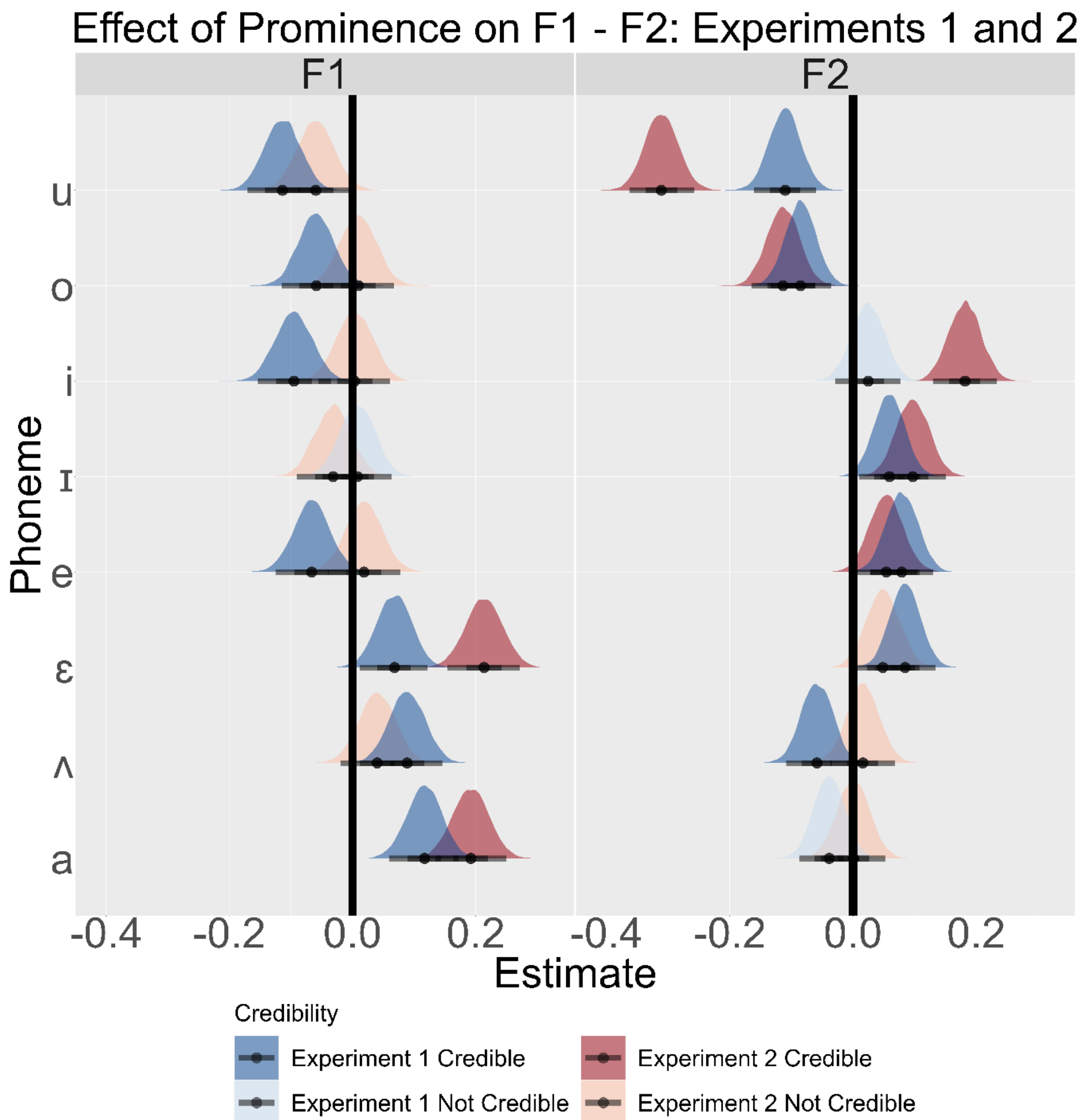
The primary goal of Study 1 has been to investigate the effects of prominence on the segmental phonetic features of vowels: How these effects manifest, how they relate to a vowel's phonological and phonetic features, and how they relate to suprasegmental prosodic features. Both experiments in the study did so by investigating the phonetic correlates of corrective focus. I originally hypothesized three possible strategies for the expression of prominence on segmental features: sonority expansion, vowel intrinsic contrast enhancement, and vowel extrinsic contrast enhancement. Sonority expansion predicts that vowels would phonetically lower under prominence; vowel intrinsic contrast enhancement predicts that vowels will shift in accordance with their phonological features; and vowel extrinsic enhancement predicts that vowels will shift away from confusable neighbors. I argue that the combined results of both experiments suggest that the effect of focus is more readily explained as vowel intrinsic contrast enhancement.

Evidence for sonority expansion, i.e., lowering of vowels for which lowering would not have resulted in phonetic peripheralization or featural enhancement, was limited to one vowel, /u/, in one sentential context in only one experiment. Likewise, the phonetic shifts observed under prominence did not reliably move vowels away from their nearest phonetic competitors, providing evidence against the vowel extrinsic hypothesis. The observed effects of prominence were related to vowels' phonological features, although this relationship did not take the form of consistent behavior across all vowels with shared phonological features. Instead, vowels that did pattern together formed natural classes such that a vowel's behavior under prominence was still predicted by its phonological specification.

Between-Experiments Comparison

The results of both experiments converge towards a characterization of prominence as being marked by phonetic enhancement in accordance with vowels' phonological features. Qualitatively, not all vowels which showed a credible effect of prominence in Experiment 1 did so in Experiment 2. Figure 14 plots the predicted effects of focus from both experiments for those vowel categories that are shared between the two experiments ({i, ɪ, ɛ, e, u, o, a, ʌ}). In the case of Experiment 2, these effects represent only those of the "object" or phrase-final context and reflect the relationship between vowels in the "confirm," condition and the "object-prominent" condition. Both graphs depict the 95% confidence interval. Among the most noticeable differences between the two experiments is the relatively few vowels showing vertical movement (F1) in Experiment 2. In Experiment 1, all plotted vowels except /ɪ/ either lowered or raised. Of the vowels that moved vertically, /u/, /o/, /ɛ/ and /ʌ/ moved along F2 as well. In Experiment 2, in contrast, vowels tended to move only along one dimension or another with the majority moving only along F2. Two other noticeable differences are in the behavior of /i/, which only moved along F1 in Experiment 1 but along F2 in Experiment 2, and that of /ʌ/, which moved along both dimensions in Experiment 1 but along neither in Experiment 2. Assuming that vowels in both experiments represent the same phonological categories, the observation that the same vowel category could behave differently in Experiments 1 and 2 provides strong evidence that prominence strategies are not strictly tied to phonological features.

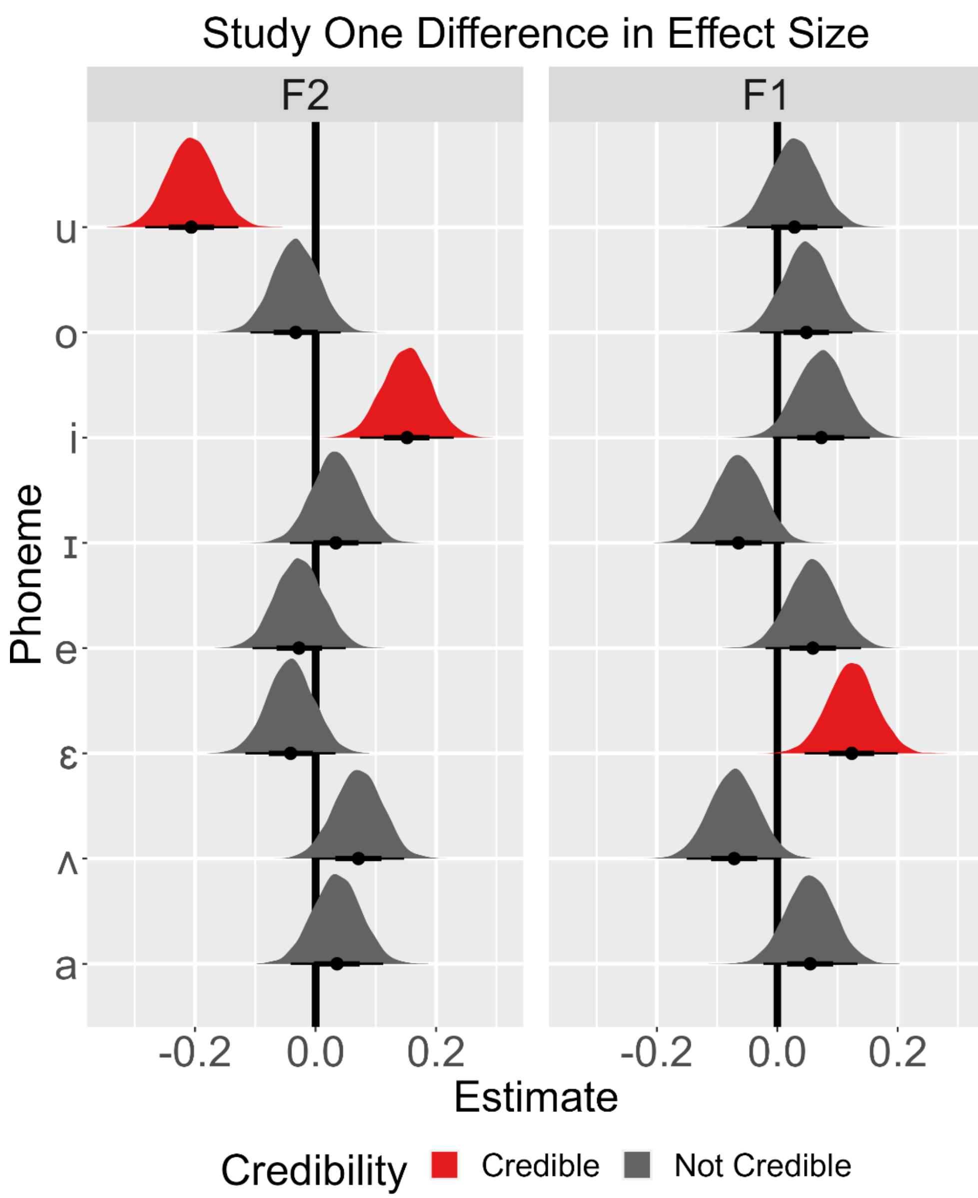
Figure 13: Effects of prominence on F1 and F2 of vowels shared between Experiments 1 and 2



However, while such qualitative comparisons between Experiments 1 and 2 suggest significant differences in how vowels behaved between experiments, a more detailed quantitative study reveals that relatively few of these observed between-experiment differences led to credible differences in vowel behavior. An additional Bayesian linear mixed-effects model was run analyzing the aggregated results of Experiments 1 and 2. Only those vowels that appeared in both experiments were included in the model. In the case of Experiment 2, only words in the object position in either non-prominent or object-prominent conditions were included. Fixed effects were included for vowel, trial type, and experiment with random effects by subject for all fixed effects.

Figure 14 shows the difference in the observed effect of prominence between Experiments 1 and 2 for each vowel. Red distributions represent credible effects. In this case, a credible effect means that the vowel in question shifted to a greater degree in one experiment than the other. As can be seen in the case of F1, /ɛ/ was the sole vowel to shift in credibly different ways in each experiment by lowering to a greater degree in Experiment 2. This is despite the fact that the effect of prominence on F1 only reached credibility in Experiment 1 for the vowels /u/, /o/, /i/, /e/, and /ʌ/. This was despite the observation that /ɛ/ and /ʌ/ differed in the credibility of their effects between experiments. In the case of F2, /u/ was observed to credibly back to a greater degree in Experiment 1 while /i/ fronted to a greater degree in Experiment 1.

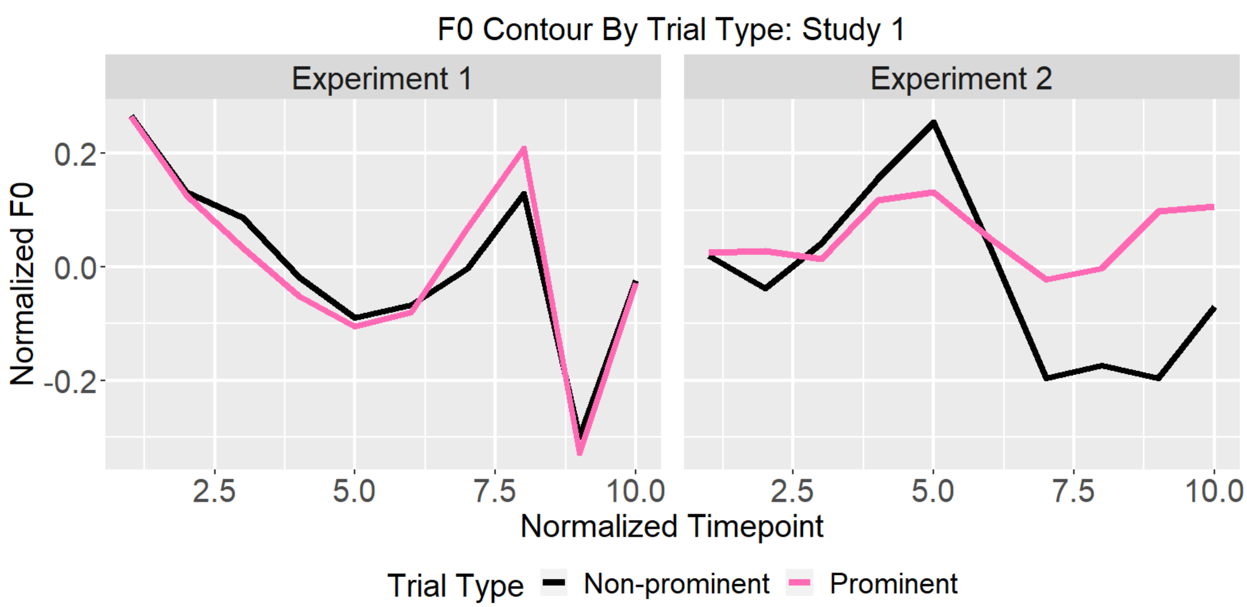
Figure 14: Difference in effect of prominence between Experiments 1 and 2



Suprasegmental Dimensions

In Chapter 2, I hypothesized that vowels which made less use of F1 and F2 to mark prominence would make greater use of suprasegmental dimensions like F0 and duration. Results do not support the hypothesis that the use of suprasegmental and segmental phonetic cues varies between vowel categories in this way. There are, however, notable differences in the use of suprasegmental dimensions between experiments. Only Experiment 2 showed credible differences between midpoint F0 based on prominence condition. Figure 15 illustrates average time normalized pitch trajectories over entire utterances in both experiments. These graphs mirror the quantitative results. No statistical difference was found between focus conditions in Experiment 1. In contrast, the F0 trajectories for object-prominent and confirm conditions of Experiment 2 differ markedly from one another and from the F0 trajectories observed in Experiment 1. Because F0 is the primary correlate of differences in intonational categories in English, these differences between experiments suggest that focus may have been prosodically marked using qualitatively different prosodic structures in each experiment.

Figure 15: Average F0 across utterances in Experiments 1 and 2



Conclusion

The goal of this dissertation is to better understand the ways in which speakers navigate trade-offs in the use of the phonetic signal to achieve multiple communicative goals. The aim of Study 1 is to better understand the use of the phonetic signal for the specific goal of signaling prosodic prominence. This serves as a necessary prerequisite for understanding its interaction with the use of the phonetic signal to further separate communicative goals: The communication of social and indexical information. While it was previously suggested that the communication of social information could make use of the same phonetic features as the phonetic marking of prosodic structure, this study provides a more thorough understanding of how they may be expected to do so by providing an in-depth empirical account of the effects of prominence on the segmental phonetic features of vowels. On a more immediate level, the results of study one provide a necessary baseline for predicting and interpreting the results of the studies to follow, presented in Chapters 3 and 4. The very limited evidence of credible differences in the behavior of vowels in each experiment establishes a degree of external validity to the following experiments by providing evidence that prominence is marked in similar ways in artificial and natural language contexts. Armed with the knowledge of what sorts of phonetic processes mark prosodic prominence, it is now possible to establish in what ways these phonetic effects may contradict or reinforce those of social signaling.

Chapter 5: Prominence and Intracategory Variation

Study 2 investigates how the phonetic encoding of prosodic prominence interacts with stylistic, socially oriented phonetic variation. I ask what happens when two stylistic variants differ from one another along a phonetic axis that also differentiates prominent and non-prominent forms. As discussed in Study 1, speakers may prioritize the clear signaling of prominence or social information at the expense of the other or may contextually alter the expression of both. Because prominence is associated with phonetic peripheralization, prioritizing its clear communication may limit a speaker's available stylistic choices to more peripheral variants. Conversely, prioritizing the production of socially preferred forms regardless of prosodic context may weaken phonetic cues to prominence. The two experiments of Study 2 ask whether speakers pursue either of these options or, alternatively, whether they adjust the communication of both social information and prominence in order to effectively communicate both simultaneously.

In order to understand how the phonetic encoding of social information interacts with that of prominence, it is critical to account for the direction, dimension, and extent to which social information and prominence exert their phonetic effects. Experiments 1 and 2 established that prominence is generally associated with phonetic peripheralization. The types of phonetic shifts used to signal social information are more varied. As discussed in Chapter 1, social information can be encoded through centralization or peripheralization along either F1 or F2. Because prominence and social information may interact differently when their effects reinforce one another versus when they detract from one another, study two is designed to examine both situations.

Experiment 3: Introduction

One obstacle to understanding how speakers balance the communication of prominence and social information is the fact that socially differentiated variants are also necessarily phonetically distinct. Because of this, any differences in how prominence is expressed on different variants could be attributable to a speakers' communicative goals *or* may simply result from differences in the phonetic targets associated with those variants. Consider for example a situation in which a speaker's repertoire includes a relatively low variant of /ɛ/ associated with Mainstream American English and a relatively phonetically high variant associated with Southern American English. The MAE and SAE vowels may move in opposite directions, i.e., further raising of the raised variant and further lowering of the lower variant, under prominence in order to communicate a social message, or as a form of phonetic enhancement, or both. The question thus becomes how to ensure that any observed differences between how different variants are made prominent are attributable to a speaker's social context and not simply the result of the phonetic target associated with a variant.

Experiment 3 seeks to answer this question by investigating how the effects of prominence differ between phonetically distinct social-stylistic variants in the absence of explicit social goals to privilege one variant over others. I do so by asking three specific research questions that correspond to three necessary prerequisites for answering this question within the specific experimental context of this dissertation. The first, RQ1, is simply whether or not speakers are able to maintain multiple phonetically distinct variants in production in the context of an artificial language. As discussed in Chapter 1, language users have been shown to be able to link social information to linguistic information within the span of an artificial language experiment. However, as previously stated, such work has not investigated phonetic variation in production in an artificial language context. It has therefore not been demonstrated that exposure to variation in speakers' input in an artificial language will lead to variation in their own productions. The short length of the experiment compared to the time it normally takes to acquire a variant may preclude speakers from accomplishing this, while the relatively socially impoverished nature of the experiment may prevent them from associating social information with a given variant.

The second question, RQ2, asks whether the distinction between phonetically differentiated variants is maintained under prominence. There are multiple potential reasons that the distinction between stylistic variants could weaken or collapse in prominent contexts even if they are maintained in non-prominent contexts. This collapse may take the form of fewer socially distinct categories being observed in prominent contexts, or in the phonetic differences between variants being diminished under prominence. This relates to the core question of this dissertation as to whether prominence constrains speakers' access to their sociolinguistic repertoire. The answer to this question has implications for interpreting to what extent any phonetic differences observed between variants in Experiment 4reflect a speaker's intended style. If participants are found to differentiate prominent forms of each variant despite not being given explicit social reasons to prefer one form over another, it would suggest that overt social preference for one form is not necessary for speakers to encode prominence differently for different variants. Conversely, if participants are not found to mark prominence differently for different variants, it would suggest that any differences observed between prominent forms in Experiment 4are likely attributable to speakers' social goals.

Finally, RQ3 asks how the phonetic alterations observed under prominence relate to the phonetic values of variants within a category. Regardless of whether or not the distinction between variants is maintained in prominent positions, it is possible for prominent variants to peripheralize according to multiple criteria. In this experiment, I ask whether prominent vowels peripheralize relative to a vowel's *average* phonetic value, over both variants or whether they do so with regard to a vowel's *most extreme variant*. Peripheralizing relative to the average vowel presents a potential problem in that a significant number of non-prominent productions may exhibit more extreme phonetic values than those of a typical prominent production. This would in turn weaken the reliability of F1 and F2 as cues to prominence. Peripheralizing relative to a vowel's entire distribution avoids this problem but could potentially weaken the distinction between prominent productions of one variant and non-prominent productions of more peripheral variants.

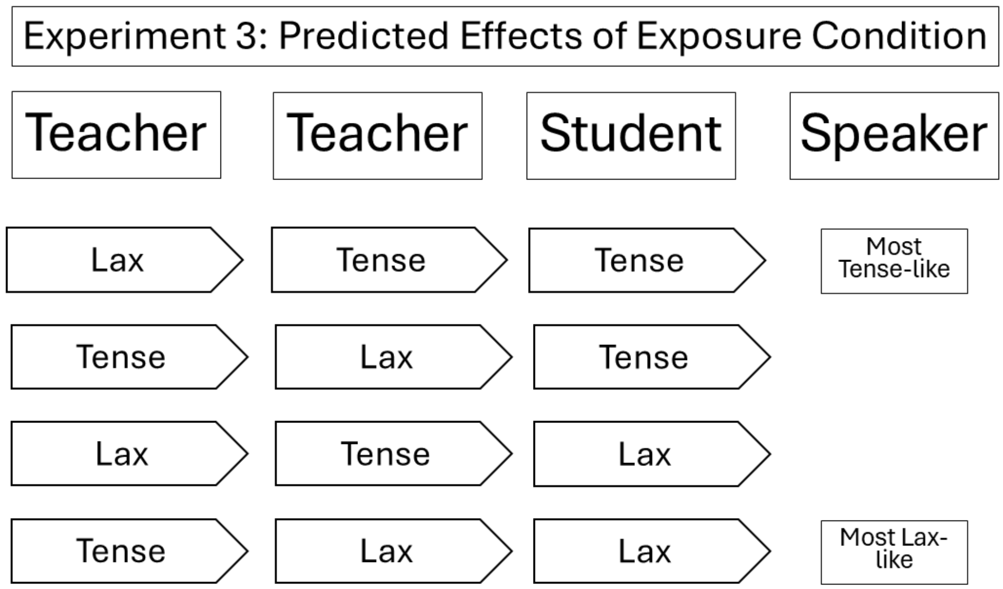
To answer these questions, Experiment 3 involved exposing speakers to an artificial language in which productions could belong to one of two “dialects.” As was the case in Experiment 2, Experiment 3 took the form of an artificial language learning study. Participants’ goal was to learn an artificial language and assist a virtual learner. Experiment 3 differs from Experiment 2 in that learners were told there were “multiple dialects,” of the language they were learning and that they may hear different pronunciations of the same word. Words of different dialects differed in the pronunciation of vowels. Speakers of dialect A or the *tense* dialect produced {i, e, a, u} in contexts where speakers of dialect B, the or *lax* dialect, produced {ɪ, ɛ, ʌ, o} respectively. These composite vowel categories are referred to as “I,” “E,” “A,” and “OU” throughout the remainder of this section. The two corresponding English vowel categories that make up each are referred to as “constituent” vowel categories. Having each vowel category in the artificial language associated with phonologically distinct English vowels was done to ensure that speakers would not have articulatory difficulty achieving the phonetic target associated with each variant. Note that {o}, like Experiment 2, is included in the “lax” set despite being a phonologically tense vowel. This was done, rather than using phonologically lax {ʊ}, because it was hypothesized that the overwhelming difference in type frequency between {u} and {ʊ} might shift speakers’ production preferences towards {u} to a high degree, regardless of whether the experimental manipulation was intended to favor {u} or {ʊ}.

Exposure to these variants was accomplished through the use of an additional model talker to the three used in Experiment 2. The new talker was identified as an additional teacher of the language who spoke a different dialect. All participants are thus exposed to one tense teacher, one tense student, one lax teacher, and one lax student. The inclusion of multiple talkers was done in order to elicit intraspeaker variation between the two dialectal forms without explicitly biasing participants towards either form. This was done by exposing all participants to three model talkers on each test trial. The first two model talkers were always the two teachers with the order of teachers balanced between trials. The third talker was always one of the two students, again balanced between trials.

I hypothesized that participants would exhibit a bias towards the most recently and/or frequently heard variant on a given trial. The four possible orderings of two “teacher” talkers followed by one “student” talker create four “exposure conditions.” This is illustrated by Figure 16. This figure shows all four possible orders of model talkers arranged according to the extent to which they are expected to trigger participants to produce tense-like vowels. Participants were predicted to produce the most lax-like tokens (resembling /ɪ/, /ɛ/, /ʌ/ or /o/) when exposed to the tense teacher, followed by the lax teacher, followed by the lax student (Tense-lax-lax or TLL exposure condition). Likewise, they are expected to produce the most tense tokens (resembling /i/, /e/, /a/ and /u/) when they first hear the lax teacher, followed by the tense teacher, followed by the tense student (Lax-tense-tense or LTT). Productions following the other two orderings (LTL and TLT) are expected to fall somewhere between these two.

Because Experiment 2 found the strongest effects in object position, this experiment elicited prominence only in this position. This also served to reduce the length of the experiment, as each trial was made longer to account for multiple model talker productions. On trials eliciting prominence, participants heard the (virtual) student produce the incorrect lexical item as the object. For example, in the case of a prominent trial where the target phrase was “Word A to word B,” participants would see the words representing “Word A,” “Word C,” and “Word B” arranged on the screen as in Experiment 2 (see Figure 17). The student model talker would produce the phrase “Word A to word D.” Importantly, the participant never hears the student talker produce the correct object of the sentence, in this case, “Word B.” This means that when on prominent trials there were effectively only two possible orderings, TL and LT, reflecting the order of the tense and lax model teachers on that trial. I hypothesized that if within-phonological category distinctions are maintained under prominence, such trials would exhibit a similar recency effect to the one seen in non-prominent trials.

Figure 16: Predicted effects of exposure condition on participant productions, Experiment 3



The results of Experiments 1 and 2 allow for the formulation of relatively precise hypotheses for how each category or variant will shift under prominence. The fact that both /ɪ/ and /i/ fronted under prominence in Experiment 2 predicts that these categories will also front in this experiment: either converging to the same target under prominence or fronting to different peripheral targets. That neither/ɪ/ or /i/ raised in Experiment 2 suggests that neither will raise in Experiment 3. However, in discussing the results from Study 1, I hypothesized that these vowels failed to raise for different reasons. I argued that /i/ may have failed to raise due to already being at the periphery of the vowel space, while /ɪ/ may have failed to raise to avoid being confused with /i/. Neither of these factors necessarily holds with regard to the combined vowel category “I” in Experiment 3. Unlike /i/, the mean of this vowel category, including both the /ɪ/ and /i/ variants, is not expected to be near the periphery of the vowel space and unlike /ɪ/, there is no phonological competitor along this dimension (in the artificial language).

If the primary phonetic strategy of prominence is peripheralization, it is therefore possible that this composite vowel will raise despite neither constituent vowel doing so.

Similar predictions can be made for the other vowels. “A” is predicted to lower under prominence related to focus. The question of whether or not this vowel will also back is similar to that regarding the raising of “I.” Although neither /ʌ/ nor /a/ backed in Experiment 2, /a/ may have failed to do so for articulatory reasons (/a/ may already be at the farthest back region of the vowel space) while /ʌ/ may have failed to back to avoid confusion with /a/. Since neither condition necessarily applies to “A,” it is possible that this vowel category will back under prominence. “OU” is straightforwardly predicted to raise and back under prominence, as both constituent vowels /u/ and /o/ backed, while /u/ raised in addition. “E” is similarly predicted to front, as both constituent vowels /e/ and /ɛ/ showed some evidence for fronting in Experiments 1 and 2. Its behavior along F1 is more difficult to predict. Both constituent vowels moved along this dimension, but in opposite directions. The composite vowel may thus be reasonably expected to raise, lower, remain stationary, or, in the case that speakers maintain multiple distinct variants under prominence, move in different directions depending on the relevant variant.

Finally, this experiment also examines the suprasegmental correlates of prominence, F0 and duration. Similarly to segmental conditions, this experiment examines F0 and duration primarily to establish a baseline to serve as a point of comparison for Experiment 4. To do so, I asks whether the effect of prominence on these dimensions, like those on segmental dimensions, varies between exposure conditions. For example, either duration or F0 may strongly differentiate non-prominent productions from prominent productions in tense-favoring exposure conditions, but not in lax-favoring, or vice-versa.

Methods

Participants

One hundred and seven Northwestern students participated in the experiment in exchange for course credit. Of these, forty-nine were excluded due to not being L1 English speakers. This left fifty-eight whose data were included in this analysis. This sample of speakers was more diverse in terms of language background compared to Experiments 1 and 2. A plurality of participants (20) reported having spent the majority of their lives before the age of twelve in Illinois. All participants identified as L1 English speakers. Twenty-six reported experience with other languages. Of these, nine reported long-term exposure to or bilingualism in an additional language (two Spanish, two Korean, one Gujarati and Hindi, three Mandarin, and one Haitian Creole). The remainder reported exposure of less than ten years. Thirty-five participants identified as female, twenty as male and one as non-binary.

Stimuli

Audio stimuli included four sets of recordings from four model talkers. Three were the same sets of recordings that appeared in Experiment 2 and one was novel to this experiment. Productions from these talkers were identical in form to Experiment 2. Recordings either consisted of a single word produced in isolation or three words spoken as a short phrase (similarly to Experiment 2, these phrases were of the form "Word A to word B"). There were eight unique words divided into two non-overlapping lists of four. Like Experiment 2, all words appeared in two forms, one with a tense vowel and one with the corresponding lax vowel. Unlike Experiment 2, participants heard tense and lax forms of every word to which they were exposed rather than hearing one set of words with tense vowels and another set of semantically distinct words with lax vowels. In Experiment 3, each participant encountered words from only one set of four words, each appearing in a tense and a lax form. Twenty-three participants were assigned to each set.

Visual stimuli consisted of the "animal," stimuli from Experiment 2. Orthographic representations were provided underneath all visual stimuli. Because each word could appear in one of two forms, orthographic representations were selected that were ambiguous between the two possible pronunciations. The vowel categories consisting of {i, ɪ}, {e, ɛ}, {u, o} and {a, ʌ} were represented by the orthographic forms "i," "e," "ou," and "a" respectively.

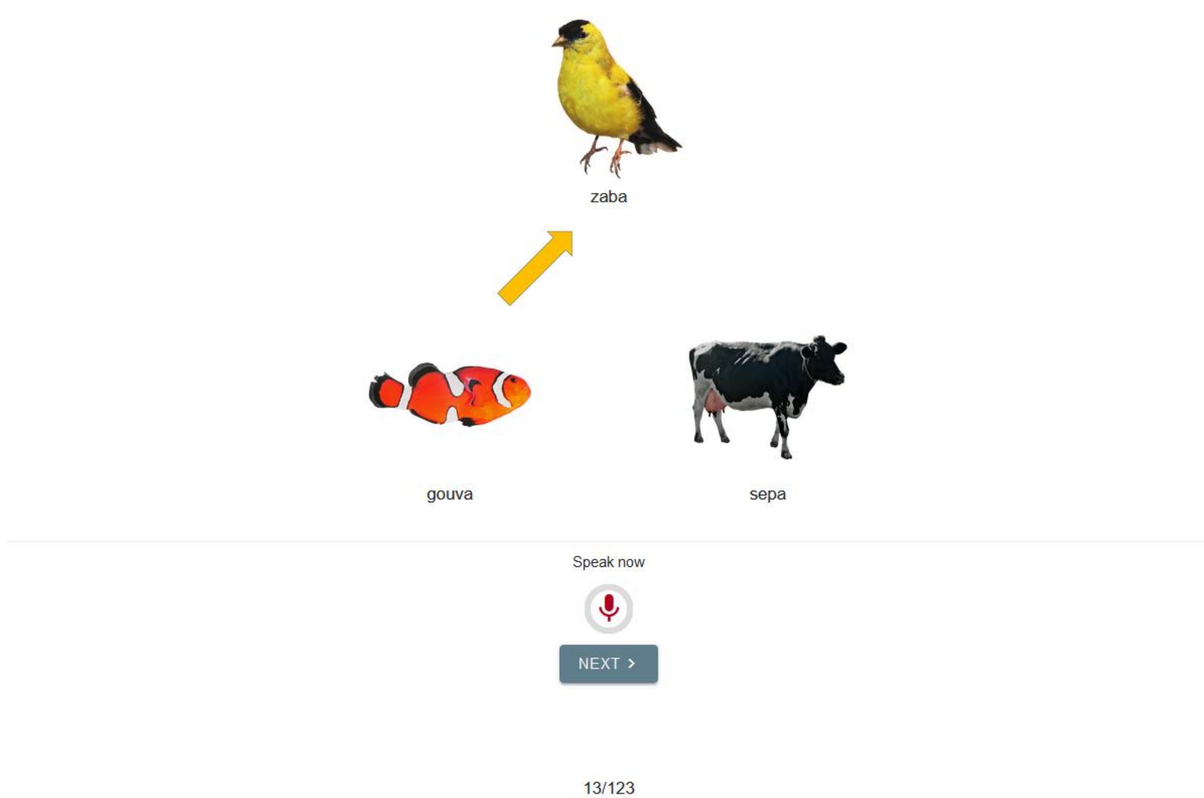
Procedure

As was the case for Experiment 2, Experiment 3 consisted of a training phase followed by a test phase. Experiment 3 was conducted in-person at the Northwestern University Phonetics Lab in a soundproof recording booth. The experiment was administered using the same custom web application as Experiment 2. After completing a consent form and background questionnaire, participants were told that they would be learning a new language. They were additionally told that the language had multiple dialects and that they may hear multiple pronunciations of the same word.

Following these instructions, participants moved on to the familiarization round containing eight trials. The screen on each trial displayed single images representing each word one at a time. Each image was accompanied by one recording from each teacher. Both teachers produced the relevant word in isolation. Participants heard the tense teacher first in half of the trials and the lax teacher on the other half. This phase was followed by a second training phase in which three images were arranged in a triangle. As described in Experiment 2, each image was associated with a target phrase of the form "Word A to Word B" where "Word A" was the word represented by the image at which the arrow started and "Word B" the word where the arrow ended. There were eight such trials in all with each word appearing twice as Word A (the subject position) and twice as Word B (the object position). The order of speakers was once again balanced between trials.

Training was followed by a testing block. The screen on each trial in the testing block displayed three images laid out in the same pattern described in the second training block. On each test round, participants first heard both teachers produce the correct phrase followed by a "guess," provided by either the tense or lax student. Participants were instructed to either confirm this guess by repeating it (non-prominent condition) or correct it (prominent condition), "using their tone of voice." Each of the four possible exposure conditions (see Figure 16) appeared three times in non-prominent and three times in prominent trials for a total of 96 trials. This included 48 confirm and 48 prominent trials. Each of the four words appeared twelve times in object position in each trial type: once in each of the four talker orderings per each of the three possible combinations of subject and object. The same was true for subject position. The experiment consisted of only a single round of training followed by testing.

Figure 17 Example of a "Non-prominent" trial



Analysis

Similar to the previous experiments, multivariate Bayesian models were created predicting F1 and F2 of test vowels from a three-way interaction between trial type (a binary predictor with levels "prominent," and "non-prominent,") vowel category (a four-way predictor with levels "I," "E," "OU," and

"A") and exposure condition (TLT, TLL, LTT, LTL). Only tokens from the object position were included in the analysis (N = 5164). Vowels were again normalized by participant and reported in standard deviation from the mean of the participant's vowel space.

RQ1 asks whether distinctions exist between variants in the non-prominent condition (i.e., from *non-prominent* trials). This was tested through a series of pair-wise comparisons between exposure conditions for each vowel. First, TLL, the exposure condition predicted to be the most lax-like, was compared to LTT, predicted to be the most tense-like. If credible differences were found between these two conditions for a given vowel, follow-up analyses were performed in which each condition was compared to the conditions that were predicted to be most phonetically similar. LTT and TLL productions of a vowel were both compared to LTL and TLT productions of that same vowel and LTL and TLT were compared with one another. A lack of a statistical distinction between two categories is taken to indicate that participants maintained a single variant in both conditions. RQ2 was similarly tested by comparing the two possible talker orderings in prominent conditions. This was done by the use of a custom coding scheme which combined TLT with TLL and LTT with LTL in prominent conditions to form TL and LT conditions respectively. Note that the maintenance of a "separate variant," in a given prominence condition only requires credible differences to occur along one of F1 or F2 in that condition.

RQ3 asked whether prominent tokens of a given category peripheralized to the point that they were more peripheral than the most *peripheral* variant of that category, or if they only peripheralized so as to be more peripheral than an *average* token of the category. For example, if "I" were to peripheralize more than an average token, but not more than the most extreme variant, a prominent "I" might have higher F2 and lower F1 than a non-prominent production of the lax variant {ɪ}, but higher F1 and lower F2 than a non-prominent {i}. It was hypothesized that, if prominent productions were in fact more extreme than an average token, but not more so than an extreme variant, it would be possible for some non-prominent tokens of the more peripheral variant to still be more peripheral than a typical prominent token. This would have the effect of reducing the reliability of F1 and F2 as cues to prominence. In the context of this experiment, the hypothesis that prominent tokens would be more peripheral than average tokens was tested by statistically comparing non-prominent productions, aggregated across all exposure conditions, to prominent productions, also averaged across exposure conditions.

Alternatively, a vowel category may peripheralize so as to become more peripheral than typical non-prominent productions of *any* of its variants, including the most peripheral variants. In the case of “I,” prominent productions would have lower F1 and higher F2 than non-prominent productions of the {i} variant. In order to test whether prominent tokens are also more peripheral than non-prominent productions of a vowel’s most peripheral variant, the average "prominent" production across exposure conditions was compared to the average non-prominent production of each exposure condition. This was done using a custom coding scheme which aggregated observations from "prominent" trials across all exposure conditions into a single “prominent,” condition. This was treated as a baseline condition. The distinction between talker orderings was maintained in confirm conditions, which were compared to the prominent baseline using treatment coding. If peripheral tokens are in fact more peripheral than the most peripheral variants, credible differences should emerge between the “prominent” conditions and all four “non-prominent” conditions. Alternatively, if the prominent condition is found not to be credibly different from one or more confirm variants, it would suggest that prominent tokens do not peripheralize to become more peripheral than all variants.

Models used to analyzed log duration and F0 were identical to those used to analyze F1 and F2. As in Experiment 2, the Praat script used to measure F0 was not able to produce measurements for all tokens included in the F1 – F2 analysis. As a result, 4928 tokens were analyzed in this model. Experiments 1 and 2 found that the effect of prominence on suprasegmental measurements were generally similar across vowel categories within an experiment, especially for F0. For this reason, F0 results are reported in aggregate across all four vowel categories.

Results

Beginning with RQ1, three of the four vowel categories exhibit evidence of credible difference between more tense-like vowels produced in the tense-favoring exposure conditions (LTT and TLT) and the more lax-like vowel produced in the lax-favoring ordering (TLL and LTL), along both formant dimensions at the 95% level, as shown in Figure 18. These are "I," (F1: $\beta = 0.252$, 95% CI = [0.173, 0.337]; F2: $\beta = -0.478$, 95% CI = [-0.558, -0.396]) "E," (F1: $\beta = -0.432$, 95% CI = [-0.514, -0.351]; F2: $\beta = 0.351$, 95% CI = [0.27, 0.429]) and "O" (F1: $\beta = 0.408$, 95% CI = [0.326, 0.488], F2: $\beta = -0.135$, 95% CI = [-0.214, -0.053]). The fourth, "A" showed no credible differences along F1 ($\beta = -0.051$, 95% CI = [-0.134, 0.028]) or F2 ($\beta = 0.076$, 95% CI = [-0.002, 0.158]). Follow-up analyses were conducted for the three vowels that showed credible effects. The orderings compared in this analysis are summarized in tables 12 and 13. For "E," and "I," all relevant orderings were differentiated across at least one phonetic dimension at the 95% level. For "OU," the two most peripheral orderings, TLT and LTT, lacked differences along F1.

Figure 18: Effect of exposure condition on non-prominent productions

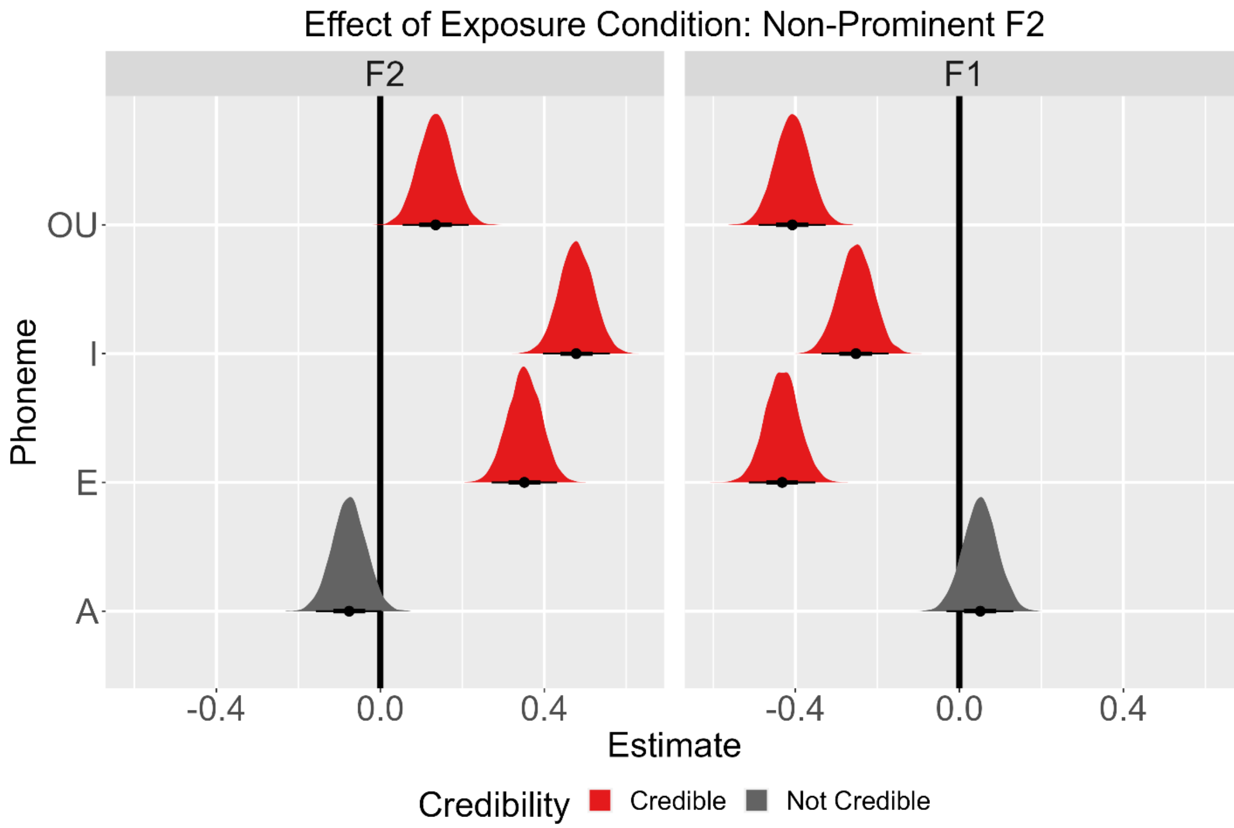


Table 12: Between-Exposure Condition Comparison: F1

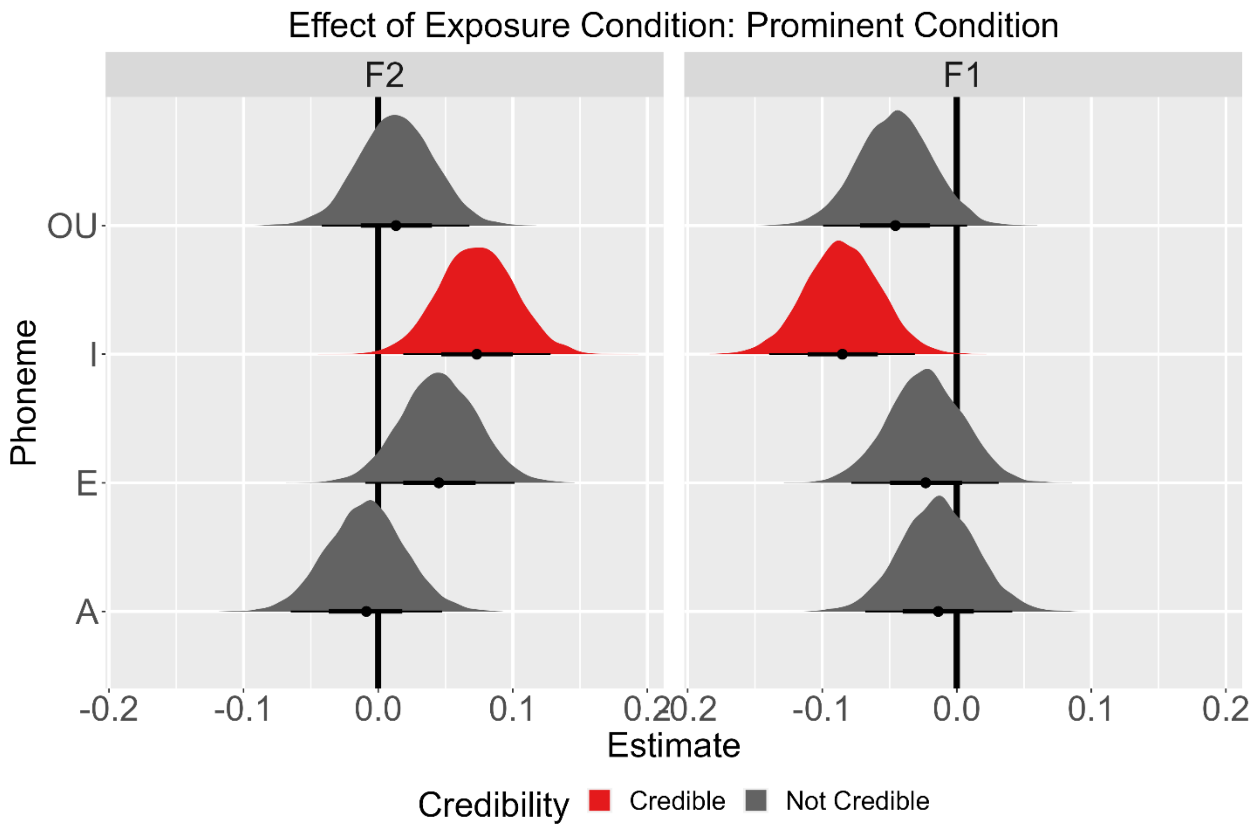
	LTT - LTL		TLL - LTL		TLT - LTT		TLT - TLL	
Vowel	**β**	**95% CI**	**β**	**95% CI**	**β**	**95% CI**	**β**	**95% CI**
E	-0.309	**(-0.385, -0.227)**	0.123	**(0.049, 0.201)**	0.072	(,-0.007, 0.148)	-0.361	**(-0.44, -0.281)**
I	-0.177	**(-0.255, -0.098)**	0.076	(-0.005, 0.154)	0.04	(-0.035, 0.117)	-0.213	**(-0.292, 0.13,)**
OU	-0.29	**(-0.37, -0.21)**	0.117	**(0.04, 0.193)**	0.041	(-0.038, 0.119)	-0.368	**(-0.447, -0.288)**

Table 13: Between-Exposure Condition Comparison: F2

	LTT - LTL		TLL - LTL		TLT - LTT		TLT – TLL	
Vowel	**β**	**95% CI**	**β**	**95% CI**	**β**	**95% CI**	**β**	**95% CI**
E	0.285	**(0.207, 0.362)**	-0.066	(,-0.146, 0.009)	-0.079	(-0.002,-0.156)	0.273	**(0.192, 0.351)**
I	0.374	**(0.297, 0.453)**	-0.104	(,-0.182, -0.019)	-0.096	**(-0.172, -0.019)**	0.383	**(0.303, 0.46)**
OU	0.132	**(0.054, 0.213)**	-0.003	(-0.081, 0.076)	-0.031	(-0.109, 0.046)	0.104	**(0.028, 0.183)**

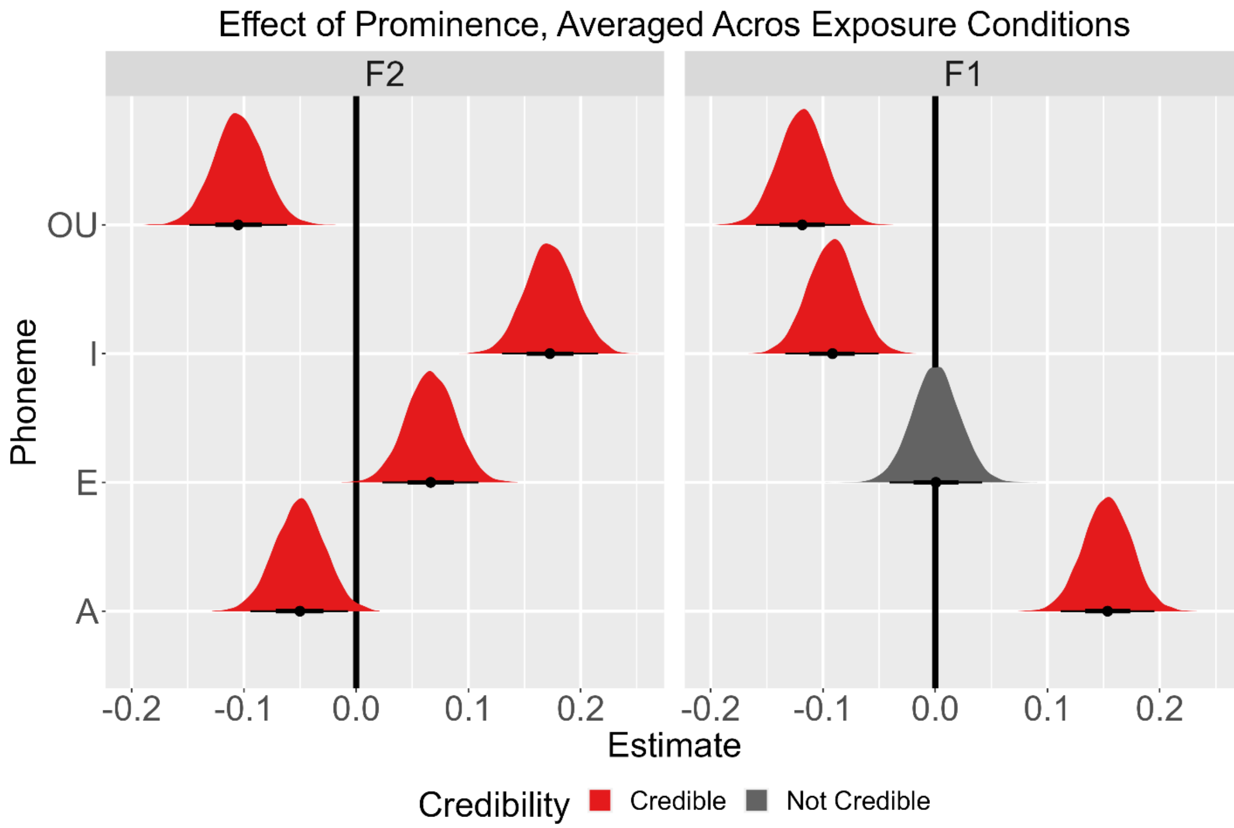
I now move on to RQ2, which asks whether the distinction between peripheralized and centralized productions, as produced under different exposure conditions, is maintained in prominent conditions. Figure 19 displays posterior estimates for the differences between these exposure conditions by Vowel. Evidence suggests that this distinction is much more weakly maintained than in prominent conditions (i.e., under corrective focus) than in non-prominent conditions. Only “I” showed credible differences between tense-favoring and lax-favoring orderings in prominent position, along both F1 (β = 0.085, 95% CI = [0.032, 0.14]) and F2 (β = -0.073, 95% CI = [-0.128, -0.018]) at the 95% level in the expected direction, with vowels in the tense-favoring order (LT) being more tense-like than vowels in the lax-favoring order (TL).

Figure 19: Effect of exposure condition on prominent productions



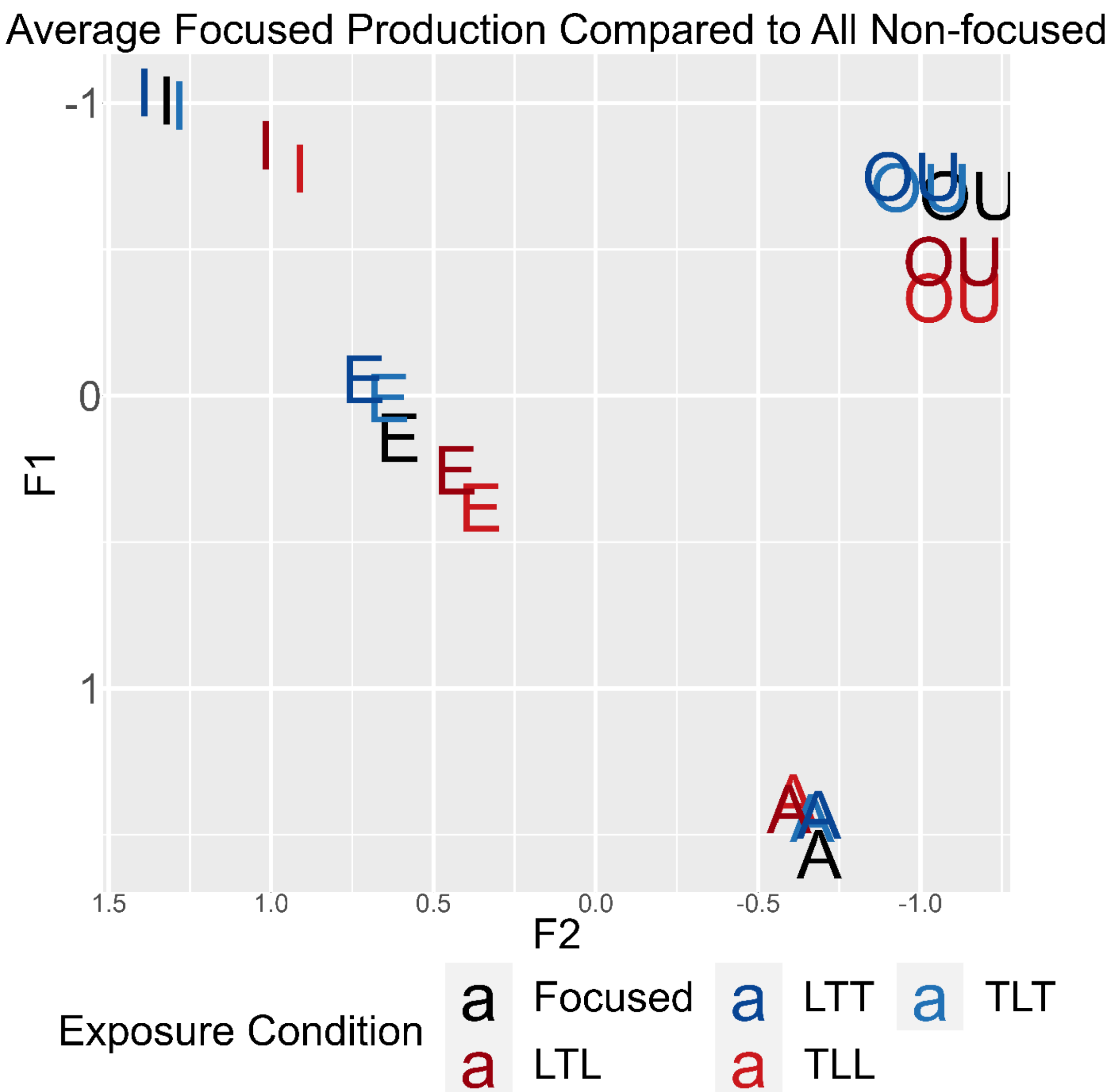
Turning now to RQ3, all four vowel categories showed robust evidence that average prominent productions were more peripheral than average non-prominent productions. This effect was found in the expected direction for all vowels and formants at the 95% level, with the exception of F1 of "E," along which prominent tokens did not credibly differ from non-prominent (Figure 20).

Figure 20: Effect of prominence, averaged across exposure conditions



Follow-up tests were performed in order to test whether prominent productions were credibly more peripheral than non-prominent vowels produced in the peripheral-favoring exposure condition, for each vowel. The predicted effect of prominence was observed only for F1 of “A”, which was more peripheral than non-prominent “A” productions in all exposure conditions. F2 of “A,” showed no credible distinction between prominent and non-prominent productions in tense-favoring favoring exposure conditions (TLT and LTT). For all other vowels and both formants, prominent productions were no more peripheral than the most *phonetically* peripheral non-prominent variant. This is illustrated in Figure 21.

Figure 21: Average prominent production compared to non-prominent productions, separated by exposure condition



In the case of F2 of "E," prominent productions were credibly backer than the most peripheral talker ordering, LTT, further front than the two most centralized LTL and TLL and not credibly distinct from intermediate TLT. F1 of "E" was unique in being the only dimension of any vowel for which prominent productions were intermediate to the most peripheral and most centralized non-prominent productions while being credibly distinct from both. F2 of "I" patterned similarly to "E," with prominent productions not being credibly distinct from non-prominent in TLT ordering and otherwise distinct in the same directions as “E.” Along F1, “I” was not distinct from either of the most peripheral talker orderings, LTT or TLT, while showing credibly lower F1 than the centralized orderings. “OU” was unique in that F2 of prominent productions being credibly further forward than the most peripheral orderings, but further back than the two centralized orderings. Finally, F1 of “OU” was found to be on average lower in prominent positions than in LTL or TLL non-prominent orderings, higher than the LTT ordering, and not distinct from the TLT ordering.

Table 14: Effect of Prominence on all Exposure Conditions: F1

Vowel	Exposure Condition	β	95% CI
A	LTL	-0.161	**(-0.226,-0.099)**
	LTT	-0.136	**(-0.198,-0.073)**
	TLL	-0.187	**(-0.252,-0.122)**
	TLT	-0.130	**(-0.191,-0.067)**
E	LTL	0.105	**(0.043,0.169)**
	LTL	-0.204	**(-0.267,-0.142)**
	LTT	0.228	**(0.164,0.295)**
	TLT	-0.132	**(-0.195,-0.073)**
I	LTL	0.151	**(0.088,0.215)**
	LTT	-0.026	(-0.091,0.034)
	TLL	0.227	**(0.161,0.295)**
	TLT	0.014	(-0.046,0.077)
OU	LTL	0.224	**(0.163,0.289)**
	LTT	-0.066	**(-0.131,-0.002)**
	TLL	0.342	**(0.278,0.407)**
	TLT	-0.025	(-0.086,0.038)

Table 15: Effect of Prominence on all Exposure Conditions: F2

Vowel	Exposure Condition	β	95% CI
A	LTL	0.091	**(0.026,0.153)**
	LTT	0.006	(-0.055,0.075)
	TLL	0.082	**(0.021,0.151)**
	TLT	0.021	(-0.042,0.083)
E	LTL	-0.172	**(-0.235,-0.11)**
	LTT	0.112	**(0.047,0.175)**
	TLL	-0.239	**(-0.304,-0.173)**
	TLT	0.033	(-0.03,0.097)
I	LTL	-0.310	**(-0.37,-0.243)**
	LTT	0.064	**(0.001,0.128)**
	TLL	-0.414	**(-0.481,-0.347)**
	TLT	-0.031	(-0.094,0.032)
OU	LTL	0.047	(-0.016,0.112)
	LTT	0.179	**(0.115,0.246)**
	TLL	0.045	(-0.021,0.109)
	TLT	0.148	**(0.082,0.209)**

Finally, regarding suprasegmental dimensions, the effect of trial type on F0 was comparable across vowels and exposure conditions (β = 0.265, 95% CI = [0.149, 0.388], averaged across exposure conditions and vowels). Figure 22 shows model estimated effects of trial type on log duration, separated by vowel and exposure condition. As can be seen, “A” was always credibly longer in prominent conditions compared to non-prominent, across all exposure conditions. “E” and “I” were credibly longer in prominent condition only in lax-favoring exposure conditions (LTL and TLL) while “U” showed the opposite pattern, being credibly longer under prominence only in tense-favoring conditions (TLT and LTT).

Figure 22: Effect of trial type on log duration, by exposure condition

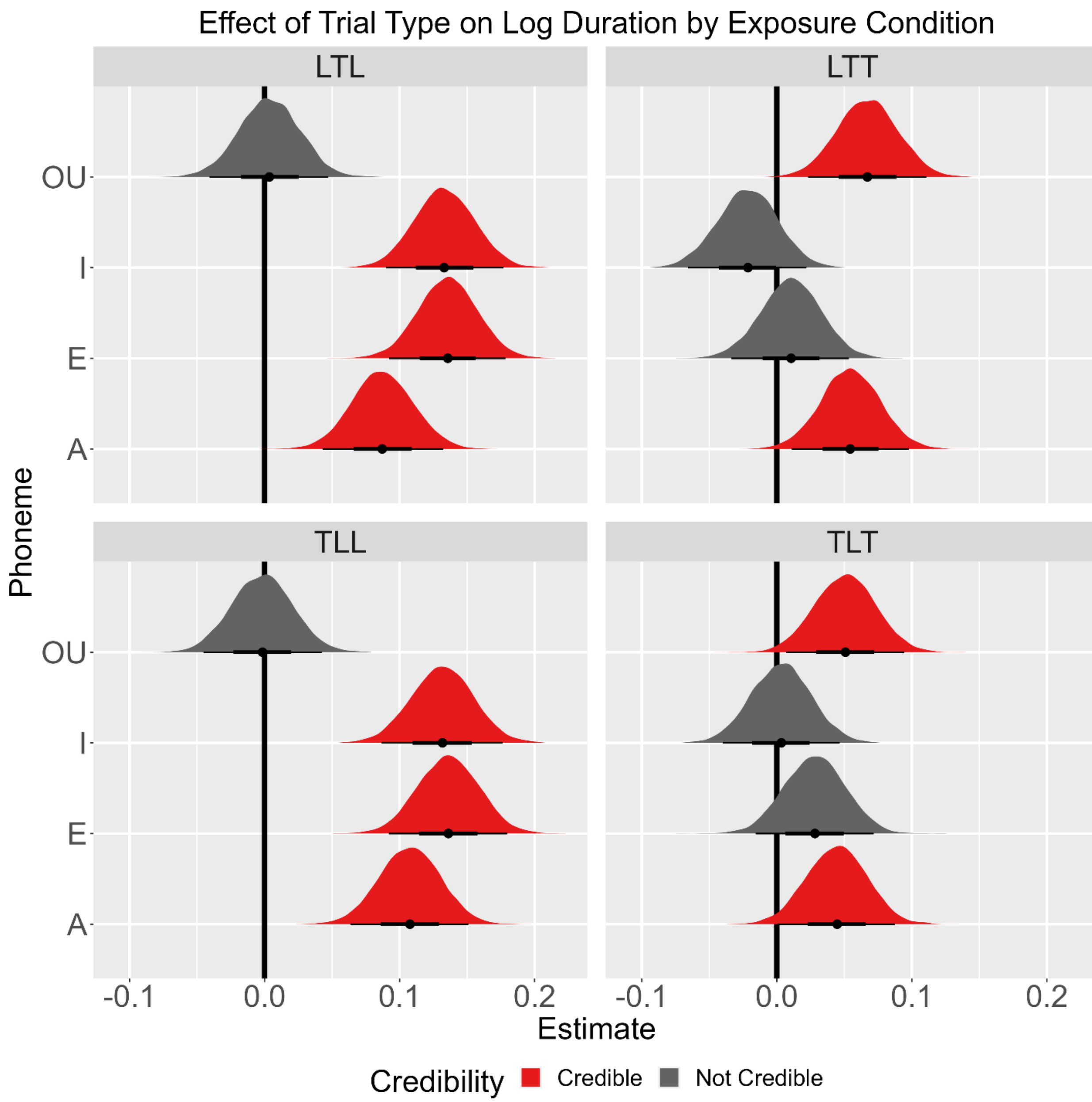


Table 16: Effect of Trial Type on Log Duration by Vowel and Exposure Condition

Vowel	Exposure Condition	Estimate	95% CI
A	LTL	0.087	**(0.044, 0.133)**
	LTT	0.054	**(0.011, 0.097)**
	TLL	0.108	**(0.064, 0.151)**
	TLT	0.045	**(0.002, 0.089)**
E	LTL	0.136	**(0.092, 0.178)**
	LTT	0.011	(-0.035, 0.052)
	TLL	0.136	**(0.092, 0.18)**
	TLT	0.028	(-0.014, 0.073)
I	LTL	0.133	**(0.092, 0.179)**
	LTT	-0.021	(-0.067, 0.02)
	TLL	0.132	**(0.086, 0.175)**
	TLT	0.003	(-0.04, 0.046)
OU	LTL	0.004	(-0.04, 0.048)
	LTT	0.067	**(0.024, 0.112)**
	TLL	-0.002	(-0.045, 0.042)
	TLT	0.051	**(0.007, 0.095)**

Discussion

This experiment sought to answer three questions. The first asked whether speakers are capable of producing two phonetically distinct forms in an artificial language, by examining vowels produced on trials that favored one form vs. the other, based on the frequency and recency to which they were exposed to each variant. One Exposure Frequency Condition favored the peripheral vowel variant by presenting that variant in two of the three exposures in the trial (frequency), and/or by presenting the tense vowel as the final exposure on the trial (recency). The other exposure condition similarly favored the lax vowel variant. The results show that this manipulation was successful in eliciting phonetically distinct productions in non-prominent (not focused) contexts in the artificial language context. Participants' productions shifted towards the most recently heard form. The second question was whether participants could extend the production of multiple forms in prominent contexts, implemented through corrective focus. While some differences were found between prominent forms based on exposure condition (favoring tense vs. lax forms), these differences were very weak compared to those in non-prominent positions. Given that Experiment 4 will bias speakers towards one of the two vowel variants, tense or lax, it seems unlikely then that this difference in prominent contexts will be stronger in that experiment.

The third question was how the phonetic and phonological properties of the composite vowel categories (combining tense and lax variants) used in this experiment relate to the types of vowel movement seen under prominence. The first major finding in this regard is that the types of movement these composite categories undergo are not strict subsets of those observed for their constituent vowels in previous studies. For instance, the composite "I" category fronted under prominence, as its constituent vowels were found to do in Experiment 2, but it also raised, which neither /ɪ/ nor /i/ were previously found to do. Similarly, while "E" fronted under prominence, as was also observed for /ɛ/ and /e/ in the Experiment 2, it neither raised nor lowered along F1. This was originally hypothesized as a possible outcome given that its constituent vowels moved in opposite directions along this dimension in previous experiments. The finding that phonetic shifting is not strictly predicted by the types of shifting undergone by constituent vowels may predict a degree of flexibility in Experiment 4. "E," for example, may reasonably be predicted to shift in either direction under social incentive.

The second major finding is that prominent vowels were often phonetically distinct from both lax and tense variants of the same vowel in the non-prominent condition. "OU" provides an example of this behavior. Prominent tokens of “OU” were on average closer to /u/ along F1 but closer to /o/ along F2. The net result of this finding was that these prominent vowels did not closely resemble either non-prominent variant. “E” provides a similar example. F1 of “E” was intermediate to the peripheral and centralized vowels. In other words, it appears to have approximated a value between these two extremes despite the fact that no such vowel was produced by the model talkers. This suggests that speakers possess a degree of flexibility in the strategies which they use to mark prominence. This provides additional evidence that speakers may utilize such distinct strategies under the social conditions of Experiment 4.

Chapter 6: Prosodic Prominence and Stylistic Variation

Experiment 4: Introduction

The previous experiments of this thesis provide evidence that speakers mark prosodic prominence on the same phonological category using different phonetic strategies**.** Experiments 1 and 2 found that speakers marked prominence differently in artificial language context compared to that of existing English words. Experiment 3 examined a situation in which speakers were required to mark prominence on novel phonological categories, each of which consisted of multiple phonetically distinct variants. It found that the strategies speakers employed to do so were distinct from those used to mark prominence in phonetically similar contexts in the absence of such salient intracategory variation. Experiment 4 investigates whether the prominence strategies speakers employ differ as speakers adapt their productions in accordance with their social situation.

Thus far, the various strategies speakers have employed to mark prominence could generally be described as involving phonetic peripheralization: Movement of phonological categories away from the center of the vowel space. Experiment 4 asks, when social contexts disfavors the use of phonetically peripheral stylistic variants, do they make use of alternative strategies for marking prominence? When the effects of social context conflict with the hypothesized peripheralizing effects of prosodic prominence, speakers may shift their productions along fewer phonetic dimensions, to lesser degrees, or in the most extreme case, in opposite directions compared to the strategies they have employed thus far in this dissertation. Experiment 4 also examines the opposite scenario, in which social context is expected to favor peripheral variants. In such situations where the demands of social information and prominence exert congruent effects, their effects may be additive. If this is the case, we would expect to see prominent productions of such peripheral variants to peripheralize to a greater degree than has been previously observed for phonetically similar vowel categories.

Sonority expansion, as discussed in Chapter 1, provides on example of an alternative strategy through which speakers may mark prominence when phonetic peripheralization is preferred by the social situation. As discussed, sonority expansion entails phonetic lowering of all vowels under prominence, regardless of their phonological features. Evidence for this strategy has been very restricted thus far. All vowels that have been found to lower under prominence have either been phonologically low as in the case of /a/ and /æ/, or phonetically below the center of the vowel space (/ɛ/ and /ʌ/}, having a higher F1 than the average of all of a speaker's vowel productions. Phonologically low vowels are also predicted to lower under the competing hypothesized prominence strategies of vowel intrinsic contrast enhancement, which proposed that vowels would shift to maximize phonetic or phonological distinctiveness; or vowel extrinsic contrast enhancement, which predicted that vowels would shift away from confusable vowels. However, if speakers modulate the phonetic strategies through which they mark prosodic prominence in accordance with their situational social needs, speakers could potentially employ sonority expansion to communicate prominence when typical strategies, such as phonetically peripheralization, did not align with their social goals.

The overall structure of Experiment 4 is identical to that of Experiment 3. Participants are exposed to the same stimuli as in Experiment 3, with identical trial structures. The experiments only differ in the instructions given to participants before exposure to the artificial language. Like Experiment 3, participants are introduced to four talkers with two model talkers producing exclusively "tense" vowels {i, e, u, a} in contexts where the other two model talkers produced "lax" vowels {ɪ, ɛ, o, ʌ}. However, unlike Experiment 3, in which two model talkers were introduced as "teachers," and other two as "students," the "teachers" were further differentiated into two roles. Prior to beginning training trials, participants were told that that one of the teachers as a "native," speaker of the language and experienced teacher, while the other was a "teaching assistant" who "does not speak the language natively, but began learning in college." The two "student" model talkers were not further differentiated and were still simply identified as "fellow students." However, as was the case in Experiment 3, one student produced exclusively tense forms while the other produced exclusively lax forms. Also like Experiment 3, "exposure condition" was manipulated within-subjects such that participants always heard three talkers on critical trials: One of the two teachers, followed by the other, followed by either the tense or the lax student.

In a between-subjects design, some participants heard the native speaker produce word forms with tense vowels (thus, the native variant), while the non-native speaker produced the same words with the corresponding lax vowel (the non-native variant). We will refer to this as the Tense Native condition, and for these participants, phonetic shifts towards the native form would be in the same direction as the shift under focus-related prominence. Other participants heard the opposite pattern, with lax vowels as the native variants and tense vowels as non-native , which we will refer to as the Lax Native condition. In the Lax Native condition, shifts towards the native variant are opposite to the expected direction of a shift under focus-related prominence. As was the case in Experiment 3, the composite vowel categories were "I" {ɪ, i}, "E," {ɛ, e}, "OU," {o, u} and "A," {ʌ, a}.

It is important to note that the primary research question of this experiment concerns the difference in prominence strategies employed for different variants. That is, the primary variable of interest is the relationship between the non-prominent form of each variant to its corresponding prominent form rather than the absolute differences between one variant and another. It is thus important to bear in mind that although the "native" status of each variant is being manipulated, the goal of the experiment is *not* to better understand the effect that manipulating "native" status has on speakers' likelihood to select on variant over another. Rather, the purpose of the nativeness manipulation is to allow a relatively high degree of certainty concerning which variant a speaker is selecting in a given context. In other words, the effectiveness of the native manipulation in biasing speakers towards a variant is taken as a working hypothesis. Importantly, it is not assumed that speakers will always shift away from non-native forms in general as there is evidence that this is not the case (Kim et al., 2011; Lewandowski & Nygaard, 2018; Wagner et al., 2021). Instead, this manipulation is expected to result in bias towards the native form due to the specific experimental context of a simulated L2 learning environment by assuming that speakers' preexisting knowledge and ideologies concerning L2 learning environments will lead them to preferentially, though not necessarily consciously, model their own speech on "native" productions.

Furthermore, while the addition of social information is meant to bias speakers' productions towards one variant at the expense of the other, there are multiple mechanisms by which such biasing could occur. These mechanisms range from overtly agentive and reflective of speaker choice to subconscious and reflective of the impact that social information may have on linguistic representation. For example, in an extreme case, a shift towards native-like forms may be largely proactive and reflect a conscious decision on the part of the speaker. It may alternatively be driven by speakers' beliefs that their "listener," either their imagined interlocutor or the experimenter, would prefer the native-like form. At the other end of the spectrum, social information may bias speakers' productions towards native-like forms indirectly by affecting speakers' cognitive representations of those variants. As proposed in Sumner et al. (2014), listeners may "socially weight" forms which they believe to be "more standard," resulting in representations of those variants being more strongly encoded in memory and thus possibly more accessible in production. The use of the term "bias" here is therefore not meant to necessarily imply a conscious or deliberate bias on the part of the speaker, but simply that productions may be "biased" towards one variant over another.

The strategies by which speakers implement prominence are expected to vary depending on which form was preferred by the situation. If participants prioritize adjusting their productions to the demands of their social context, i.e., producing vowels more similar to the socially preferred variant than the dispreferred variant, then we predict the effect of prominence of vowel production to vary between the

Tense Native and Lax Native conditions. Essentially, participants are predicted to mark prominence in ways that maximize phonetic similarity to the preferred form, as opposed to ways that would clearly convey phonological features in a typical English context. In doing so, they will be able to mark prosodic prominence without shifting towards a socially dispreferred form. Participants biased to prefer the lax vowels are predicted to peripheralize to a lesser degree or along fewer dimensions compared to participants biased to prefer the tense vowel. In the most extreme case, vowels produced in the Lax Native condition may exhibit prominence-related shifts in the opposite direction to what was observed in Experiment 3.

For example, in Experiment 3, F1 and F2 of "I" in varied between in non-prominent contexts, being more phonetically similar to whichever constituent vowel ({i, ɪ}) participants had most recently been exposed to. "I" also exhibited raising and fronting under focus-related prominence relative to non-prominent productions of the same vowel. These findings inform predictions about the behavior of "I" in Experiment 4. First, I predict that, rather than shifting their productions of non-prominent forms towards the most recently encountered form, non-prominent forms will resemble the socially preferred form. Productions are thus expected to show higher F2 and lower F1 in non-prominent productions in the Tense Native condition relative to the lax native.

Furthermore, in the Lax Native condition of Experiment 4, we may expect that speakers will avoid marking "I" as prominent in ways that phonetically shift it towards the non-native, tense form. I thus hypothesize that "I" in the Lax Native condition will exhibit only fronting, only raising, or will remain stationary, relative to non-prominent productions of the same vowel in the same Tense Native condition. The most extreme possible strategy is predicted to be similar to the Sonority Expansion hypothesis discussed earlier. In this situation, "I" would lower. This may have the effect of enhancing its similarity to the socially preferred lax vowel, resulting in prominent productions of this vowel being those that most clearly mark the situationally preferred stylistic guise. In contrast, "I" in the Tense Native condition is hypothesized to raise and front at least as much as was observed in Experiment 3, and possibly more so if the effect of the social bias is additive to that of focus-related prominence.

Similar predictions can be made for all other vowels. For the constituent vowels of “E”, {ɛ, e}, prominence effects were found to shift in opposite directions along F1 in Experiments 1 and 2 while the corresponding vowel in Experiment 3 did not vertically shift at all. Given that we have observed variable effects of prominence on these vowels, I predict that productions of the “E” vowel will raise under prominence in the Tense Native condition and lower in the Lax Native condition. Its movement along F2 is more straightforward. /e/ was found to front under prominence in Experiments 1 and 2, while /ɛ/ only fronted in Experiment 1, where it did so to a lesser extent. Productions of “E” in the Tense Native condition are therefore expected to front at least as much as productions of “E” in the Lax Native condition. As was the case for “I,” productions of “E” in non-prominent contexts are expected to show higher F2 and lower F1 in the Tense Native condition, compared to the Lax Native.

The constituent vowels of “OU”, {o, u} differ from those of “E” in that, rather than the tense vowel being both higher and more horizontally peripheral than the other, previous experiments have found that F2 of {u} is generally higher than that of {o}. Non-prominent productions of “OU” are therefore expected to have lower F2, but also lower F1, in the Tense Native condition. The constituent vowels of “OU” have also both been found to exhibit lower F2 under prominence. “OU” would thus be predicted to back in both conditions without compromising a speaker’s intended social message. However, it would only be expected to phonetically raise under prominence in the Tense Native condition, as doing so in the Lax Native condition would shift it closer to the dispreferred {u}.

Finally, “A,” {ʌ, a} backed and lowered under prominence in Experiment 3. Therefore, in the Tense Native condition, where the preferred variant is {a}, productions of “A” are predicted to be further back and lower compared to the non-prominent productions of the same vowels. Predictions for how prominence will affect productions of “A” in the Lax Native condition, where {ʌ} is the preferred variant of the “A” category, are less clear. Unlike the other vowel categories, a sonority enhancing lowering of {ʌ} under prominence would shift productions of that the vowel towards, rather than away from, the non-preferred variant, {a}. It is thus predicted that productions in the Tense Native condition will either remain stationary or centralize.

An alternative strategy by which speakers could potentially mark prominence when their typical strategies for doing so run counter to the variant preferred by their current social situation is by doing so exclusively through suprasegmental means. When speakers are socially biased against more peripheral forms, they may not shift their productions along F1 and F2 under prominence at all, but may instead produce greater distinctions between prominent and non-prominent productions along F0 and duration. In the case of this experiment, this would predict that speakers in the Lax Native condition will make use of larger differences in F0 and duration to mark prominence than will speakers in the Tense Native condition.

Finally, given the results of Experiment 3, it is possible that exposure condition will interact with social preference and prosodic prominence. I hypothesize that speakers’ productions may still shift in the direction of the most recently heard form, regardless of which form is socially preferred. I predict that this biasing effect will be weaker than the one observed in Experiment 3, such that productions will overall still more closely resemble the socially preferred form than the most recently encountered form. I predict, however, that, all else being equal, productions will still more closely resemble more recently encountered forms.

Methods

Participants

Sixty-one participants were recruited through the Northwestern linguistics subject pool. Twenty-four were excluded due to not being L1 English speakers. Data from the thirty-seven remaining speakers are presented here. Stimuli and procedure were identical to Experiment 3. The only difference was that the initial instructions introducing the teachers introduced one of the teachers as a "native speaker and experienced teacher" and the other as a "teaching assistant who began learning the language in college." Twenty participants were in the Lax Native condition and seventeen in Tense Native. As in the previous experiments, participants were further divided into subgroups exposed to non-overlapping sets of vocabulary items, Set A ({vupə, vopə}, {dazə, dʌzə}, {kebə, kɛbə}, {zifə, zɪfə}), and Set B ({guva, gova}, { zabə, zʌbə},{sepə, sɛpə} {pibə, pɪbə }). Seventeen participants took part in the Tense Native condition, ten of whom heard Set A and seven set B. Twenty participants in the Lax Native condition heard Set A and ten Set B.

Procedure

Procedure was identical to Experiment 3, except for the addition of additional biographical information in the instruction section. In addition, on test trials, participants were presented with a written description at the bottom of the screen which either read "Teacher," "Teaching Assistant," or "Student," depending on which model talker was speaking.

Analysis

As in previous experiments, participants' productions were normalized before analysis. Normalized values are thus reported in standard deviations from the mean of a speaker's vowel space. However, because this normalization technique scales the vowel space and recenters it at zero, it is problematic for analyzing the between-subjects manipulation of this experiment. This is illustrated by Figures 23 and 24. Figure 23 shows mean unnormalized F1 and F2 values for both Tense Native and Lax Native conditions; Figure 24 shows normalized measurements. Both figures show averages across both trial types. As can be seen from these figures, despite the fact that the Tense Native and Lax Native conditions showed clear differences in unnormalized F1 and F2, these differences were not always observable in normalized formant values. This is especially visible for the peripheral vowels "I," "OU," and "A."

Figure 23: Mean Unnormalized F1 and F2 values of both experimental conditions. Means are averaged across exposure conditions and trial type

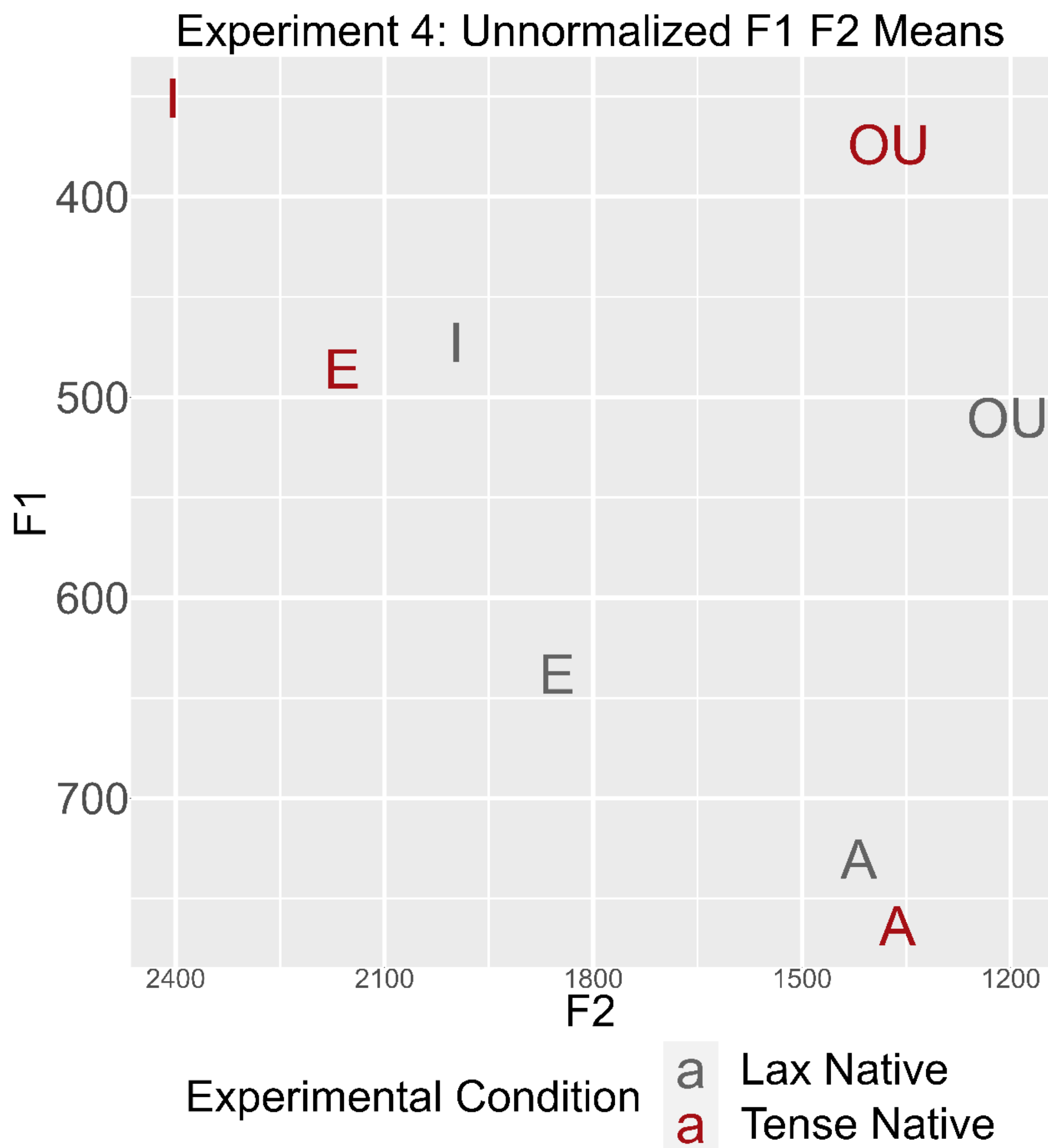
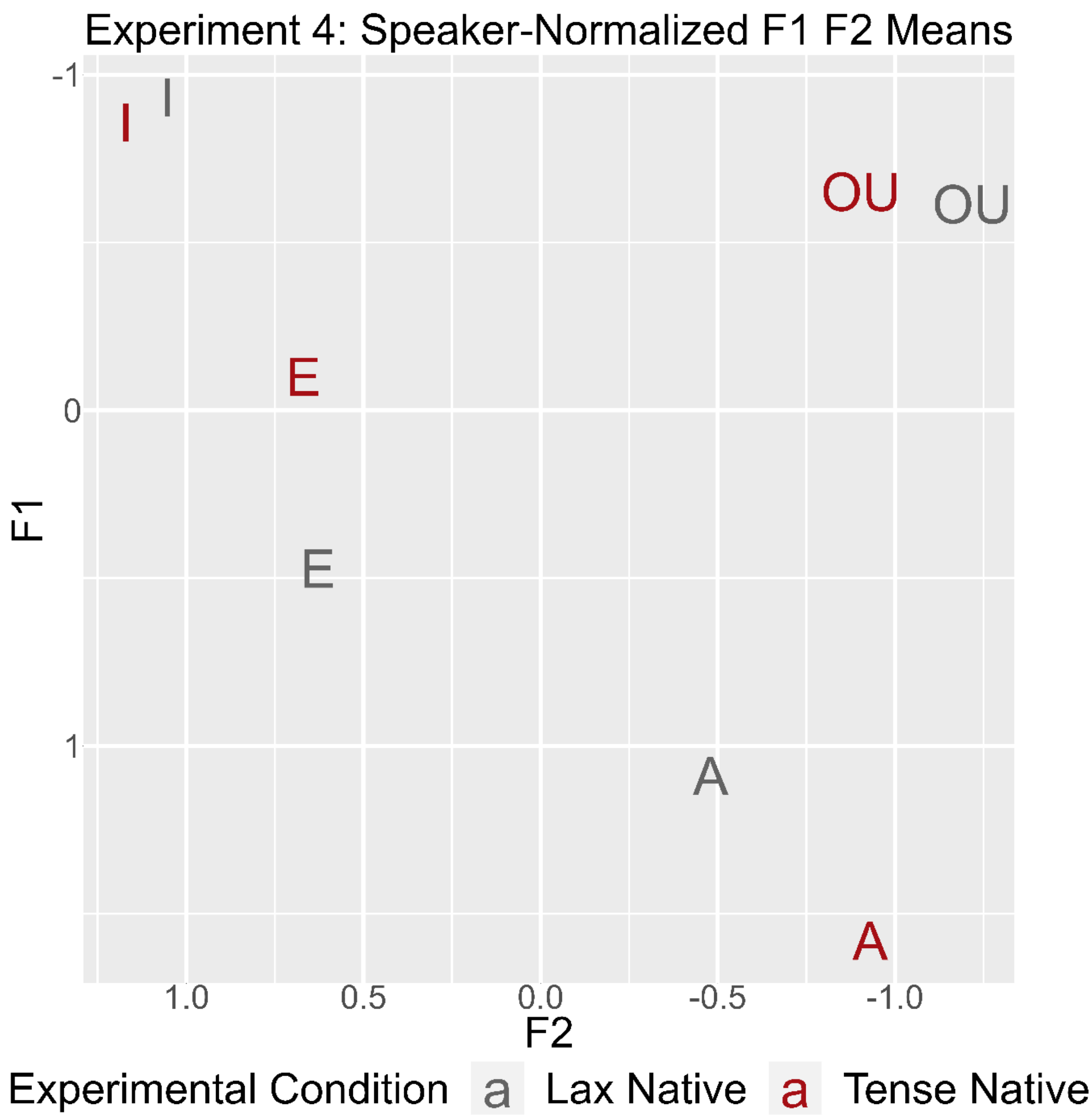


Figure 24: Mean normalized F1 and F2 values of both experimental conditions. Means are averaged across exposure conditions and trial type



F1 and F2 were thus analyzed using two models: One for normalized and one for unnormalized values F1 and F2 values. These models were identically structured. Fixed effects included vowel category (a four-way variable with levels “I,” “E”, “A” and “OU.”) Experimental Condition (*Lax Native* and *Tense Native*) and trial type (“Prominent” and “Non-prominent”), as well as the three-way interaction between these. Exposure condition (a four-way categorical variable as in Experiment 3 with levels TTL, TLT, LTT, and LTL) was also included, as well as its interactions with vowel, Experimental Condition, and trial type Condition. Random slopes and intercepts were included for vowel, experimental condition, trial type by speaker. Identically structured models were created analyzing F0 and log duration.

Results

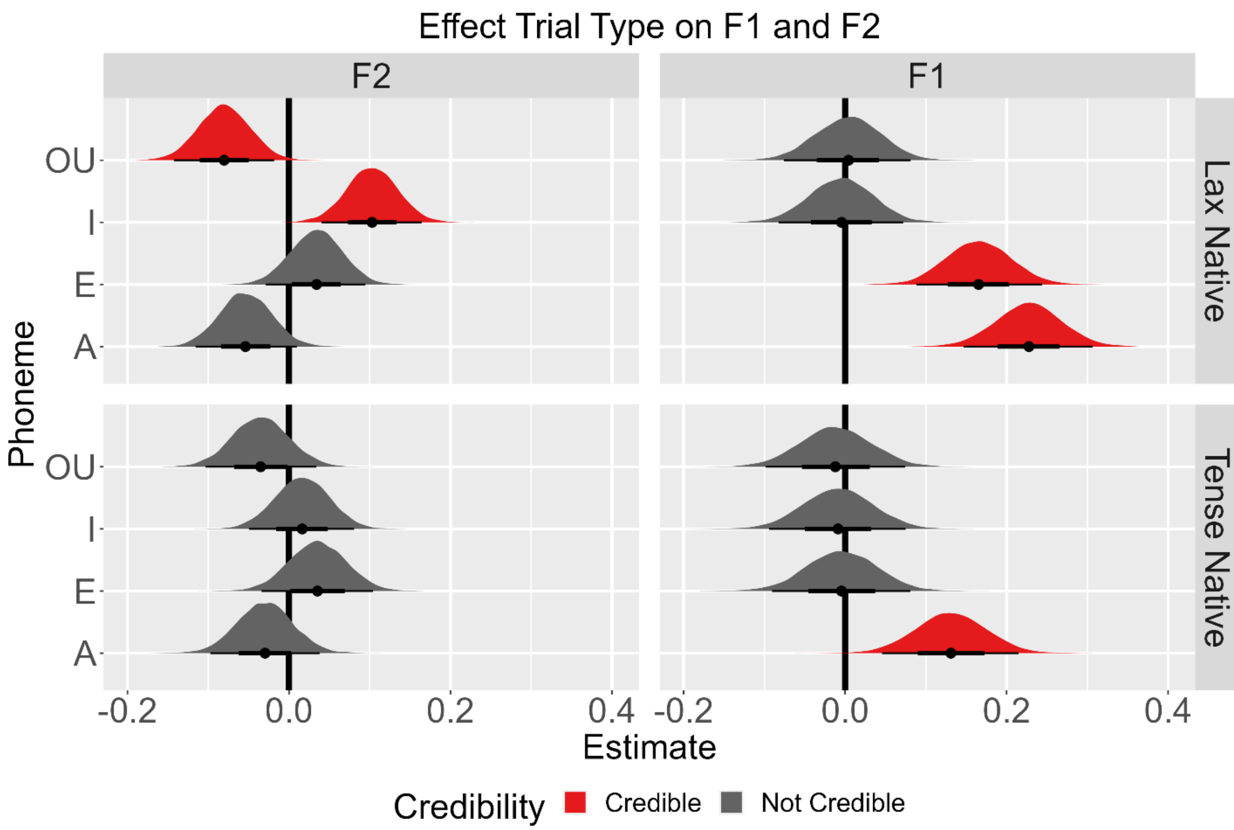
As previously illustrated in Figures 23 and 24, unnormalized F1 and F2 measurements, averaged across trial types and exposure conditions, demonstrate a clear effect of experimental condition for all vowels except "A." In the case of "A," no credible differences emerged between the Tense and Lax Native conditions. "I," "E," and "OU" produced in the Tense Native condition exhibited lower F1 than corresponding vowels produced in the Lax Native condition. The front vowels "E" and "I" exhibited higher F2 in Tense Native condition, while the back vowel "OU" showed lower F2 in the Lax Native condition. The fact that Lax Native exhibits lower F2 than tense native is expected given the widespread fronting of /u/ that has been observed in North American English.

Table 17: Effect of Experimental Condition on F1 and F2, Averaged Over Talker Order and Trial Type

	F1		F2	
Vowel	**β**	**95% CI**	**β**	**95% CI**
A	24.493	(-27.273, 78.806)	-39.371	(-109.908, 32.544)
E	-147.518	**(-205.633, -87.052)**	243.092	**(96.772, 381.281)**
I	-120.468	**(-170.045, -63.747)**	324.536	**(158.1, 481.466)**
OU	-135.325	**(-176.363, -92.347)**	124.075	**(32.44, 239.089)**

I turn now to the effects of prominence within experimental conditions, illustrated in Figure 25. In the Tense Native condition, only "A" was found to show a credible effect of prominence, with lowering in the height dimension indicated by higher F1 values under focal prominence (β = 0.131, 95% CI = [0.041, 0.21]). In the Lax Native condition, prominence effects were consistently observed with all vowel categories showing effects along at least one formant dimension. Under focal prominence, “I” fronted as indicated by higher F2 (β = 0.103, 95% CI = [0.041, 0.164]) and “OU” backed as indicated by lower F2 (β = -0.08, 95% CI = [-0.141, -0.018]) . “E” (β = 0.165, 95% CI = [0.089, 0.244]) and “A” (β = 0.228, 95% CI = [0.147, 0.307]) both lowered under prominence, as indicated by higher F1.

Figure 25: Effect of trial type on F1 and F2 of Experiment 4



I turn now to the effects of exposure condition. As Figure 26 illustrates, in the Lax Native condition, there were no differences between exposure conditions with regard to the effect of trial type on F1. “A” and “E” always exhibited higher F1 in prominent trials under all exposure conditions, while “I” and “OU” never showed a credible effect of prominence under any condition. Effects were more variable with regard to F2. “A” showed a backing effect of trial type only in the lax-favoring exposure conditions (LTL and TLL). “E” showed the opposite pattern, showing credible degree of fronting the prominent Condition relative to the non-prominent only in tense-favoring conditions (TLT and LTT). “I” patterned similarly to “E” along F2, fronting most strongly in TLT and LTT, but also showed a smaller fronting effect in TLL. “OU” patterned with “A,” with a credible backing effect of prominence only in LTL and TLL.

Figure 26: Effect of trial type by exposure condition, Lax Native

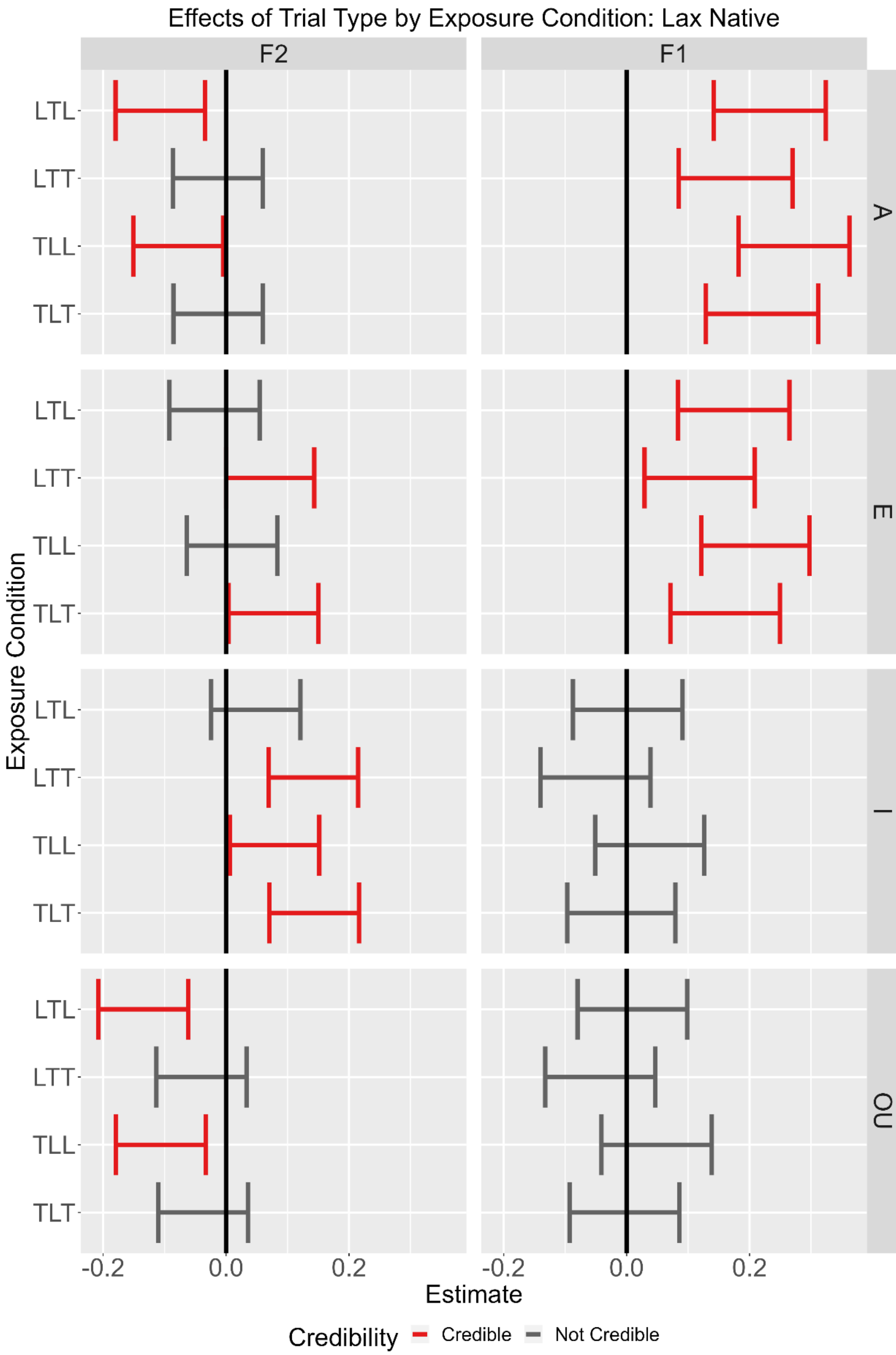
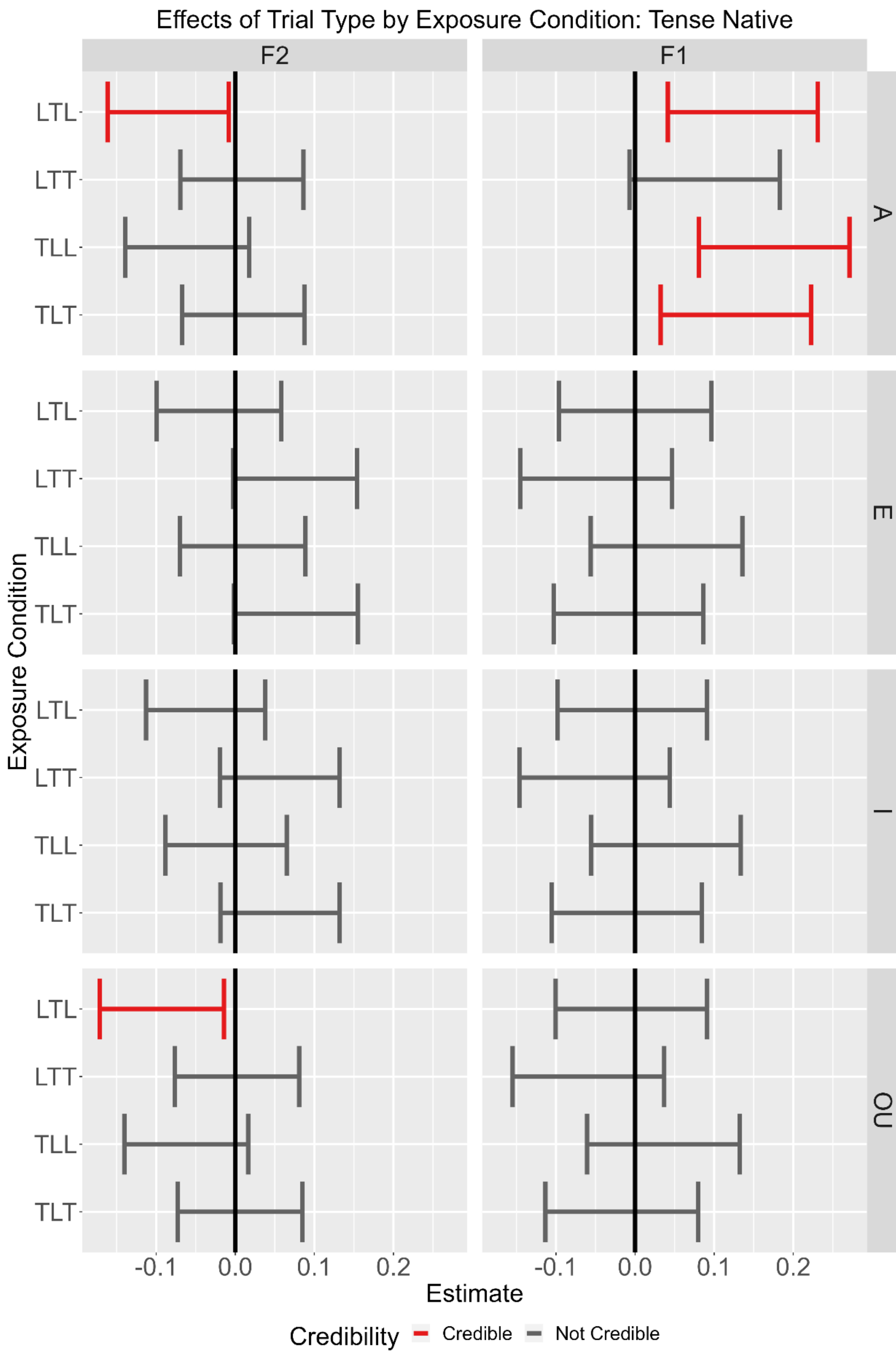


Figure 27 illustrates the effect of trial type by vowel and exposure condition in Tense Native condition. As can be seen, the effects of prominent condition is more consistent between vowels and exposure conditions in this condition. Almost all combinations of exposure condition and vowel showed no effect of trial type on F1 or F2. The exceptions were “A,” which showed a credible lowering effect of prominence in all conditions except LTT and a credible backing effect in LTL; and “OU,” which also showed a credible backing effect in LTL.

Figure 27: Effect of trial type by exposure condition, Tense Native



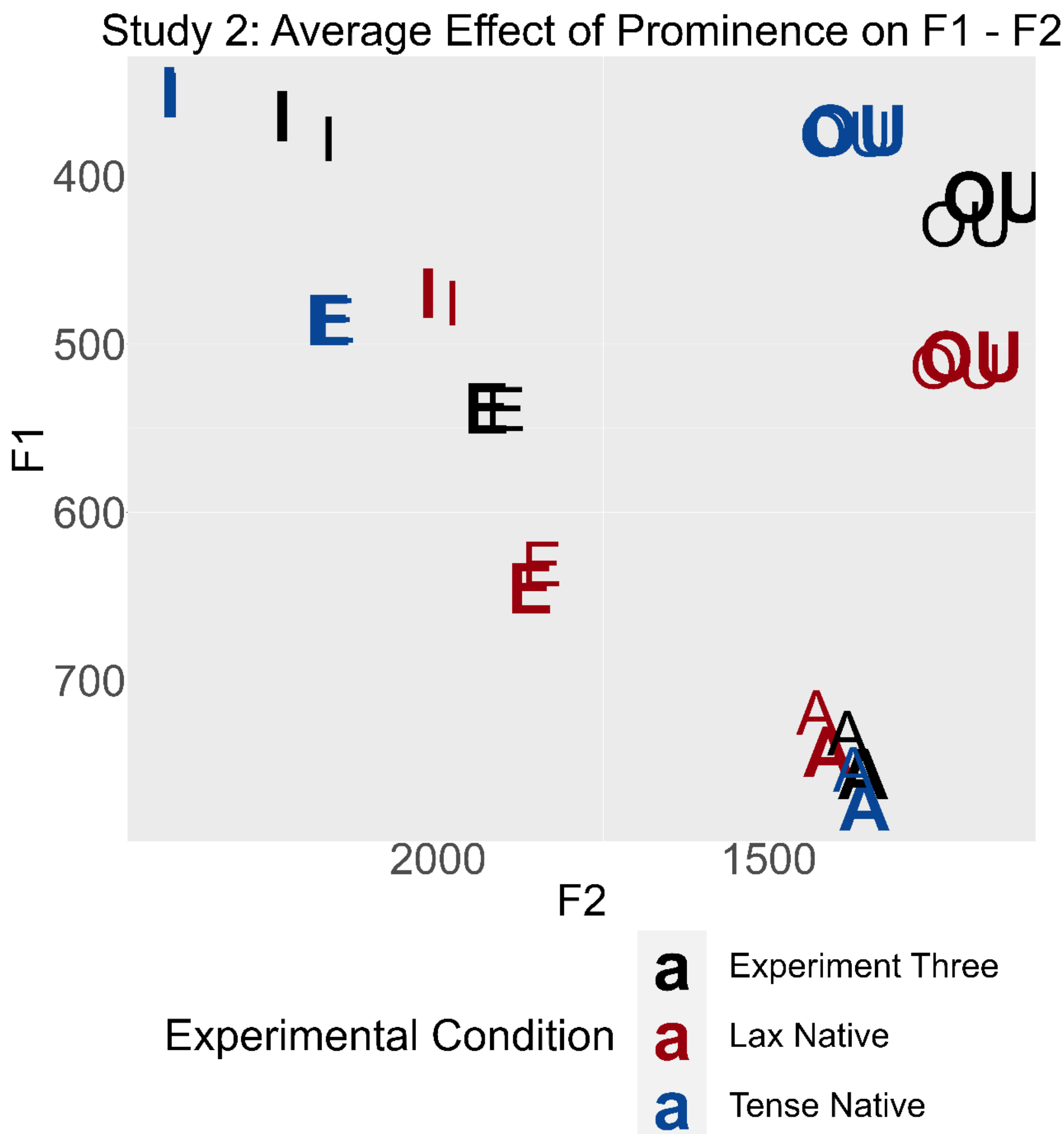
Finally, suprasegmentals are reported aggregated across vowels and exposure conditions. Trial type was found to be a credible predictor of F0 only in Lax Native condition, where midpoint F0 was found to be higher in prominent conditions (($\beta = 0.15$, 95% CI = [0.083, 0.217])). The same was true of duration, with prominent productions credibly longer than non-prominent ($\beta = 0.036$, 95% CI = [0.013, 0.058])

Chapter 7: Study 2 Overall Discussion

To summarize the results of Experiment 4, participants produced phonetically distinct vowels and employed different prominence strategies based on experimental condition. The prediction for this experiment was that speakers biased to prefer "lax" variants {ɪ, ɛ, ʌ, o} would shift along fewer dimensions, to a lesser degree, or in opposite directions to those biased towards "tense" productions (i, e, a, u}. Contrary to these predictions, participants in the Lax Native condition made more consistent use of F1 – F2 movement to mark prominence than those in Tense Native. Furthermore, when the effects of prominence and social preference incentivized movement in opposite directions**,** specifically, when the social bias was for the centralized variant (the Lax Native condition), participants still employed phonetic peripheralization as a prominence strategy despite the fact that doing so sometimes shifted productions towards the dispreferred peripheral variant. When social information and prominence favored the same variant, as in the Tense Native condition, evidence was only found for prominence effects on a single vowel category, "A." The absence of prominence effects for the other vowel categories indicates that F1 and F2 were substantially weakened as overall cues to prominence. In other words, in the Tense Native condition, prominent and non-prominent vowels were no longer reliably differentiated along these two formant dimensions.

In both the Tense Native and Lax Native conditions of Experiment 4, speakers employed distinct strategies for marking prominence which aligned with their social-communicative goals. Experiment 3 presented a different pattern of prominence effects, where all vowels except "E" peripheralized along both F1 and F2. Figure 28 compares Experiments 3 and 4 in the average raw (not normalized) F1 – F2 values for data aggregated over prominent conditions. It becomes clear that participants in the Tense Native condition of Experiment 4 produced on average vowels that were more peripheral than those in Experiment 3. By producing vowels as maximally peripheral regardless of prominence condition, participants in the Tense Native condition effectively maximized the alignment of their vowels with those produced by the Native speaker at the expense of using F1 – F2 space to differentiate prominent and non-prominent productions. In contrast, participants in the Lax Native condition produced vowels that were more centralized than those in Experiment 3, and also never used both F1 and F2 to mark prominence on the same vowel. In other words, in Experiment 4, the observed shifting effects of prominence on vowels in the Lax Native condition were generally consistent with a strategy of using the F1 and F2 dimensions to mark prominence while minimizing movement towards the tense forms that were socially dispreferred in that experimental condition.

Figure 28: Mean F1 – F2 values of all vowel categories in prominent and non-prominent contexts across Experiments 3 and 4. Prominent conditions are in larger font and bolded



Looking at vowels in the Lax Native condition of Experiment 4, “OU,” provides the clearest example of the restricted peripheralizing effect of prominence. In shifting this vowel to a more back position under prominence (lower F2 values), participants were able to simultaneously mark “OU” productions as more prominent while shifting them away from the socially dispreferred {u} variant. However, unlike Experiment 2, where {u} productions raised under prominence, “OU” did not raise in Experiment 4, not even in the Tense Native condition where {u} was the socially preferred variant. Given that the sole difference between Experiments 3 and 4 was the presence of a social bias, it is reasonable to conclude that participants in the Lax Native condition of Experiment 4 avoided raising as a strategy to mark prominence with "OU" to avoid shifting productions of "OU" towards the dispreferred {u} variant.

The situation is similar with regard to "E," which was found to lower in the Lax Native condition of Experiment 4. This strategy allowed speakers to mark prominence on "E" in a way that also shifted its production away from the dispreferred form. However, unlike productions of "OU" in the same condition (Lax Native), we do not find prominence effects on "E" in the same formant dimensions as in Experiment 3. In Experiment 3, "E" fronted, but did not move vertically under prominence, while in the Lax Native condition of Experiment 4, "E" lowered with no concomitant horizontal movement. By employing this strategy, speakers were able to mark prominence in a way that aligned with the strategy that Experiment 2 found for the phonetically similar English vowel /ɛ/, while also shifting their productions away from the dispreferred {e} form.

The case of the {ɪ} variant in the Lax Native condition was different in that, by employing the strategy of peripheralizing along F2 that was seen for the phonetically similar English vowel /ɪ/, speakers also shifted their productions *towards* rather than away from the dispreferred tense variant. This shifting towards the dispreferred variant would have also happened, arguably to an even greater degree, had speakers employed the same strategy of phonetic raising and fronting observed for the "I" class in Experiment 3. speaker However, by peripheralizing along only F2 in the Lax Native condition of Experiment 4, rather than along both F1 and F2 as in Experiment 3, speakers were able to mark prominence while mitigating, if not entirely avoiding, movement towards {i}. Finally, while the social manipulation of Experiment 4 was successful in eliciting between-conditions differences among "I," "OU," and "E," this was not the case for "A." This category exhibited comparable phonetic realizations and similar prominence strategies in Experiment 3 and in both conditions of Experiment 4.

A more in-depth quantitative comparison between Experiments 3 and 4 slightly complicates this picture. Figures 29 and 30 illustrate the results of a follow-up aggregate model examining Experiments 3 and 4. The model included predictors for trial type (prominent or non-prominent), experimental condition (Experiment 3, Experiment 4 Tense Native, Experiment 4 Lax Native), vowel, and talker order. Results presented have been marginalized over talker order. In these figures, red distributions represent effects that were stronger in Experiment 3 than the corresponding condition in Experiment 4. As Figure 29 illustrates, in the case of F1, the observed effect of trial type in Experiment 3 was stronger than the corresponding effect in Experiment 4 for all vowels and conditions except for "E" and "A" in the Tense Native condition. This provides evidence the speakers were in fact making use of this dimension to a lesser degree in the Lax Native condition than in Experiment 3. It also shows that this was true for "I" and "OU" in the Tense Native condition as well. This accords with the observation that these vowels did not credibly raise in either condition in Experiment 4, but did so in Experiment 3.

For F2, as seen in Figure 30, the sole credible difference in effect size was found for "I," where the effect in Experiment 3 was found to be stronger than in Tense Native. This adds an important caveat in the interpretation of Experiment 4. Namely, although the effect of prominence on F2 reached credibility for "I" and "OU" in Lax Native but not in Tense Native, the effects observed in both conditions were not credibly different from those in Experiment 3. It is therefore not the case that the vowels which moved in Lax Native along this dimension did so to a lesser degree than in Experiment 3.

Figure 29: Difference between effect of prominence between Experiment 3 and both conditions of Experiment 4 on F1

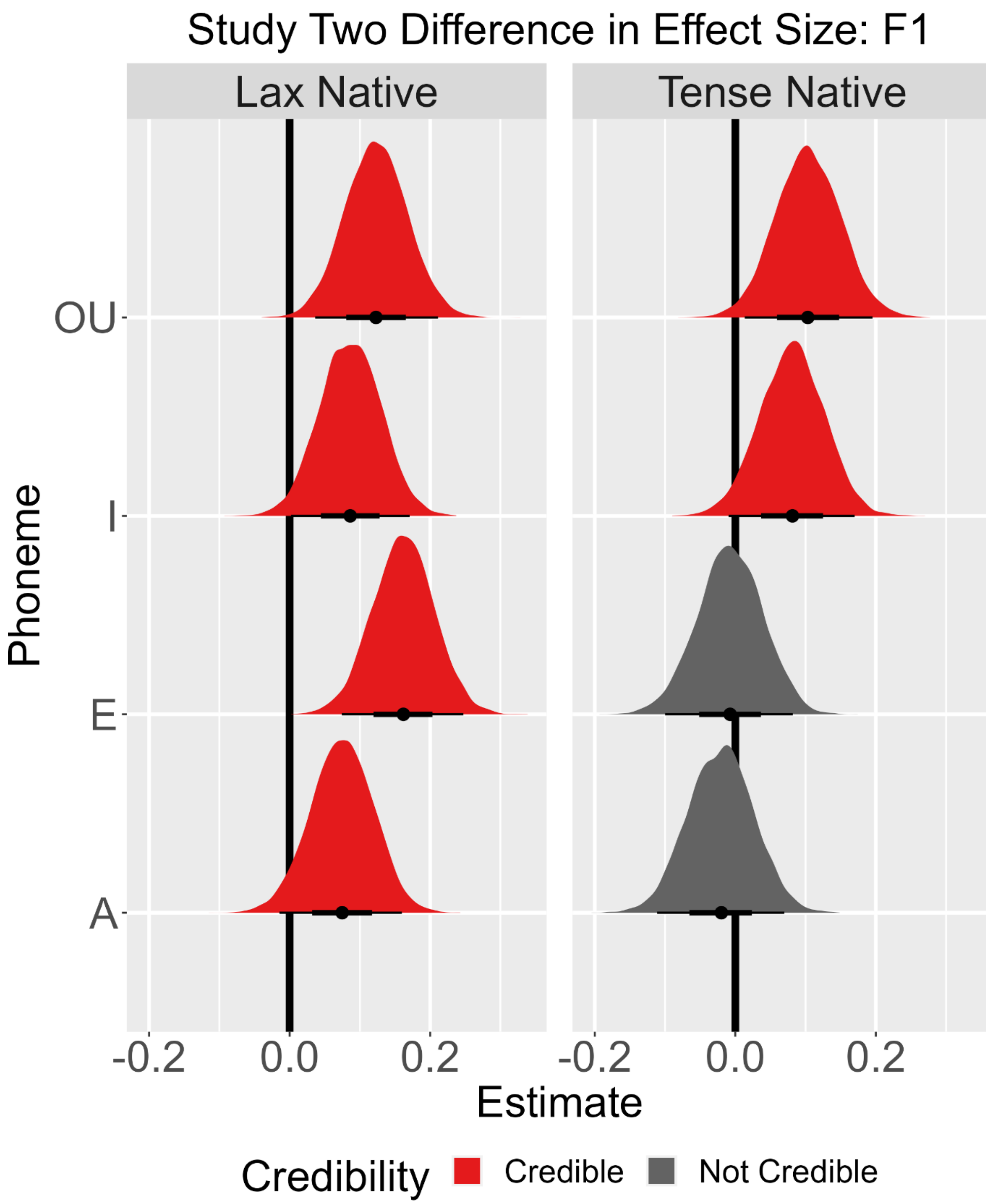
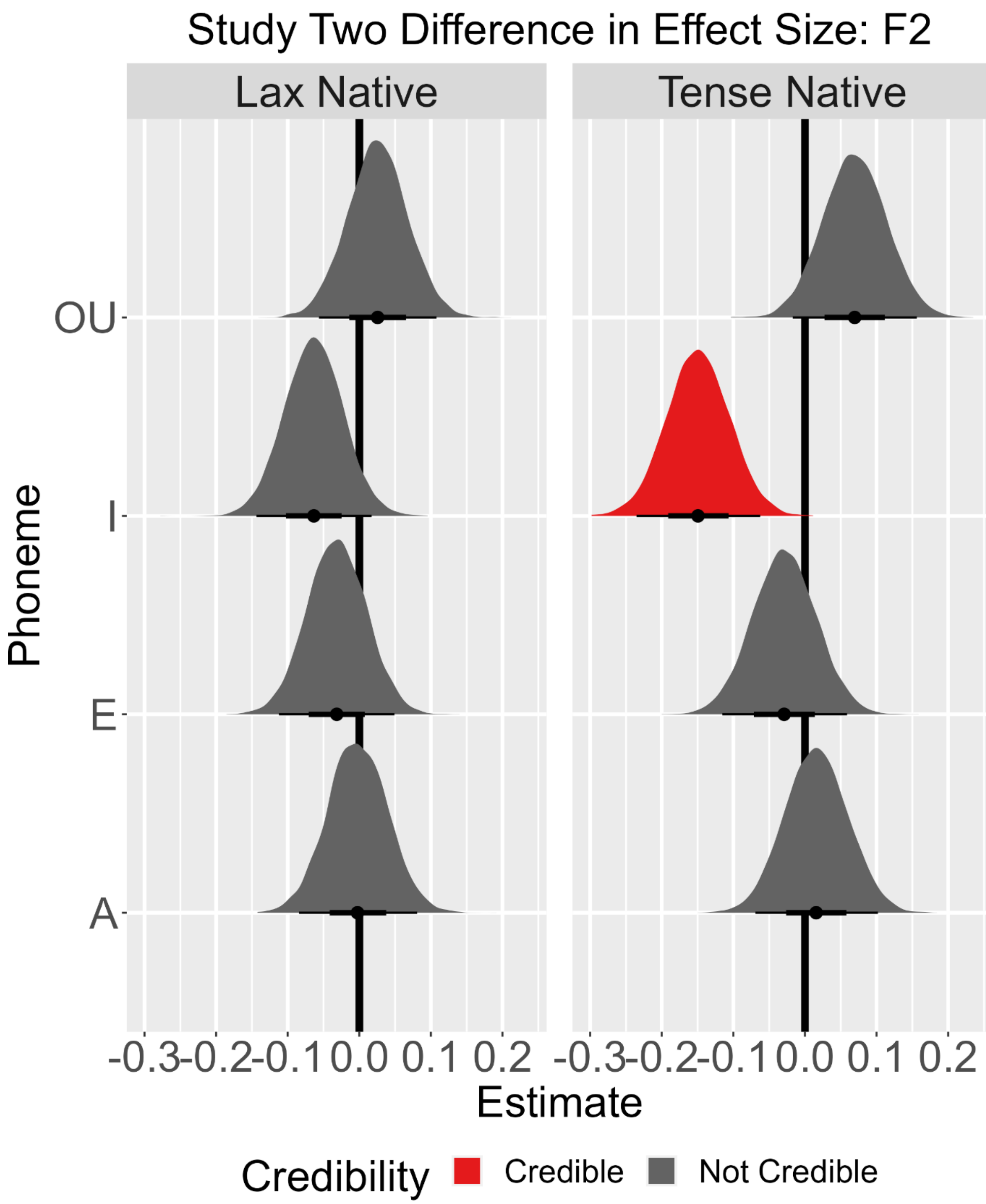


Figure 30: Difference between effect of prominence between Experiment 3 and both conditions of Experiment 4 on F2



Because prominence was predicted to be less strongly marked along segmental dimensions in the Lax Native condition, it was predicted to be more strongly marked along F0 and duration in this condition. While it was not found to be the case that speakers in the Lax Native condition relied less on segmental dimensions, they did, somewhat surprisingly, appear to make stronger use of suprasegmental cues to prominence. These suprasegmental results are somewhat surprising. They suggest that, in the Tense Native condition, prominence was not reliably marked along any of the four phonetic dimensions examined: F1, F2, F0, or log duration. The lack of a credible distinction along F1 and F2 could potentially be explained as participants consistently converging towards the preferred variant regardless of prominence condition. The same is also possible for duration as tense vowels (with the exception of /u/) were typically longer than their lax counterparts. The F0 results are somewhat more puzzling. Previous experiments did not suggest that tense vowels were significantly higher than their lax counterparts. It stands to reason then that the higher overall pitch of the Tense Native condition was related to the specific targeting of tense vowels, or explicit avoidance of lax vowels, rather than being inherent to the vowels in question. The nature of this relationship remains an area for future study.

The results of Experiment 4 suggest that speakers can maintain distinct prominence strategies for different variants, i.e., for the socially preferred tense variant or the socially preferred lax variant. Nonetheless, the between-subjects design of the experiment leaves open the question of whether an *individual* speaker can do so. Outside of a laboratory setting, an individual speaker typically has access to multiple forms that are situationally preferred in different contexts. Participants in this experiment were exclusively incentivized to prefer one form at the expense of another. Experiment 3 found that speakers collapsed the distinction between peripheral and centralized variants under prominence, but it suffers from the opposite problem in that while speakers were presented both variants to imitate, for many novel vocabulary items, they had no incentive to maintain this distinction in their own productions. It is therefore at least possible that, had speakers been incentivized to produce both variants, the phonetic distinctions between the variants would have lessened.

Overall, Experiments 3 and 4 provide evidence that social incentives can influence the phonetic strategies speakers employ to signal prominence but that these strategies are constrained by phonetic factors. In the case of these experiments, "social incentives" came in the form of salient biographical information about the appropriateness of modeling productions on one speaker over another in order to perform the specific social role of assisting in L2 language learning. Participants varied the extent to which they shifted their productions under prominence and the dimensions along which they did so depending on the form that had been identified as native. However, the direction in which prominent vowels shifted was largely consistent within vowel categories across social-communicative contexts. Evidence suggests that the direction in which a vowel category shifts under prominence is determined by its phonetic characteristics such that when prominence effects are observed in F1 x F2 space, prominent vowels shift towards the periphery of this space compared to non-prominent vowels. Therefore, if "phonetically extreme," is equated with "phonetically peripheral," it would appear that speakers are restricted to employing prominence strategies that result in more "extreme" productions.

Chapter 8: General Conclusion

This dissertation was designed to ask how speakers balanced their use of limited phonetic resources to pursue potentially conflicting communicative functions. I investigated the specific case of potential conflicts between speakers' the use of the phonetic signal to signal discourse-level pragmatic information and the use of stylistic variation to adapt their speech to a specific social context. Experiments 1 and 2 provided necessary empirical foundations for asking this question by clarifying the phonetic strategies by which prominence is communicated along the same phonetic dimensions that communicate social information. I described multiple factors that defined a phonetic strategy by asking along which dimension, to what extent, and in which direction does a given vowel shift under prominence. I also asked whether the answers to these questions depended on a vowel's phonetic and phonological features. I compared evidence for three such strategies: Sonority expansion, vowel intrinsic contrast enhancement, and vowel extrinsic contrast enhancement. Sonority Expansion predicted that prominent vowels would phonetically lower (exhibit higher F1) regardless of identity. Vowel intrinsic contrast enhancement predicted that vowels would peripheralize in ways determined by their phonological characteristics, shifting in directions that maximize perceptibility of their phonological features. Finally, the vowel extrinsic contrast enhancement hypothesis predicted that vowels would enhance their phonological features so as to become less confusable with nearby phonological neighbors.

Experiment 1 investigated this question through the use of a speech production paradigm that elicited prosodically prominent vowels through association with focus. Results, I argue, did not straightforwardly align with any of the three hypothesized strategies. While some non-phonologically low vowels were found to lower, these vowels (/ɛ/ and /ʌ/) were also phonetically well below (higher F1) than the midpoint of speakers' vowel spaces. While vowels did generally peripheralize horizontally in accordance with their phonological backness features, the relationship between vowel height and phonetic shifting was less straightforward. Vowels which shared phonological height features did not necessarily pattern together. For example, /e/ raised while /ɛ/ lowered, and phonologically high /ɪ/ and /ʊ/ failed to raise despite mid vowels /e/ and /o/ doing so. There was also little evidence that nearby phonological neighbors blocked phonetic shifting under prominence, as predicted by the vowel extrinsic contrast enhancement hypothesis. Regardless of the lack of a consistent, phonologically driven strategy, these findings provided novel understanding of the behavior of the vowels of Mainstream American English under prominence. They also allowed me to form more precise hypotheses for the following studies.

Experiment 2 translated these results into the context of the artificial language paradigm upon which the rest of this dissertation was based. Comparisons between the results of Experiments 1 and 2 found that vowels tended to behave similarly in both experiments, with no credible differences in the effect of prominence on a vowel category based on its experimental context. This provides evidence that the artificial language paradigm used in the following experiments could successfully elicit prominence patterns comparable to those speakers use in their native language. There were, however, a handful of examples of the phonological categories examined in the setting of familiar English words in Experiment 1 being marked as prominent differently in the context of unfamiliar nonce words. /i/ and /u/ both shifted horizontally to a greater extent in Experiment 2 than in Experiment 1, and /ɛ/ lowered to a greater extent than in Experiment 1. The fact that vowels behaved differently under prominence in Experiment 2, despite being phonologically identical to the corresponding vowels in Experiment 1, provides strong evidence that a although a vowel's phonological characteristics may strongly influence its behavior under prominence, they do not necessarily determine such behavior. I argue that these results suggest a characterization of prominence strategies as flexible, with speakers adopting different strategies in different contexts.

Experiment 3 showed that even in the absence of explicit social motivation to prefer one form over another, the existence of socially linked variation enabled speakers to utilize novel strategies for marking prominence. Comparing the values produced in Experiments 2 and 3 shows that productions of composite vowel categories in Experiment 3, where participants were introduced to two talkers representing different dialects, were phonetically intermediate to those of corresponding constituent vowel categories in Experiment 2, where participants heard the same inventory of vowel forms which were not associated with dialectal differences. This is interesting, given that participants were never exposed to such phonetically intermediate productions. Instead of veridically reproducing specific heard productions, participants essentially averaged over phonetic input to produce novel phonetic targets.

They were then able to mark the corresponding phonological categories as prominent in ways that aligned with the findings of Experiment 1 in relying on phonetic peripheralization. Speakers employed strategies to mark prominence on these composite vowel categories that were distinct from those used by their corresponding English vowels

Experiment 4 tested the hypothesis that speakers could employ prominence strategies that shifted productions towards situationally preferred variants, rather than solely employing peripheralization. This was not found to be the case. While speakers did mark prominence differently depending on which variant was situationally preferred (in a between-subjects design), they did so by peripheralization even when doing so brought them away from preferred forms. This finding contributes to the debate surrounding the effect of phonetically “enhancing,” environments on socially motivated variation. As discussed in Chapter 1, previous work by Clopper and colleagues has argued that enhancing environments tend to bias speakers towards “standard” as opposed to “dialectal” forms. While this dissertation differs in describing variants as “native,” or “non-native,” rather than “dialectal” and “standard,” it provides evidence that prominence shifts productions in specific directions rather than towards or away from specific variants. To the extent that prominence behaves like other enhancement favoring contexts, this would predict that speakers would shift towards “standard” forms only when those forms are also more peripheral. As different sources of phonetic enhancement have been argued to result from different underlying mechanisms, this serves as a potential area for future investigation.

Overall, this dissertation advances our understanding of how speakers navigate the need to use a limited phonetic signal to balance multiple, potentially competing communicative functions. Social variation has been argued to associate most readily to aspects of language where it is unlikely to interfere with the communication of linguistic information (Eckert & Labov, 2017). This dissertation provides evidence that, in the case of the use of prosodic structure to encode pragmatic linguistic information, this is not necessarily the case. Instead of fine-grained phonetic detail functioning solely or even primarily to encode linguistic or structural information, it appeared to function chiefly as a means for speakers to situate themselves within a specific social context. The primary contribution of this dissertation is in providing evidence that the interactions between multiple sources of informative phonetic variation may not depend on a single, fixed strategy for prioritizing one source of variation over another, but rather are flexible and reflect the concrete communicative needs of the specific context in which speakers find themselves.

References

Arvaniti, A. (2020). The Phonetics of Prosody. In A. Arvaniti, *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.411

Arvaniti, A., Katsika, A., & Hu, N. (2024). Variability, overlap, and cue trading in intonation. *Language*, *100*(2), 265–307. https://doi.org/10.1353/lan.2024.a929737

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31–56.

Baker, R. E., & Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech*, *52*(4), 391–413.

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145–204.

Beltrama, A. (2020). Social meaning in semantics and pragmatics. *Language and Linguistics Compass*, *14*(9), e12398. https://doi.org/10.1111/lnc3.12398

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, *5*(5), e10773.

Büring, D. (2016). *Intonation and meaning*. Oxford University Press. https://books.google.com/books?hl=en&lr=&id=tky1DAAAQBAJ&oi=fnd&pg=PP1&dq=+B%C3%BCring,+D.+(2016)+Intonation+and+Meaning.&ots=nnscjie28t&sig=em0bxD98h2ZFxZ7XZIQ_H0U43WI

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of/ɑ, i/in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867–3878.

Cho, T. (2016). Prosodic Boundary Strengthening in the Phonetics–Prosody Interface. *Language and Linguistics Compass*, *10*(3), 120–141. https://doi.org/10.1111/lnc3.12178

Clopper, C. G., & Dossey, E. (2020). Phonetic convergence to Southern American English: Acoustics and perception. *The Journal of the Acoustical Society of America*, *147*(1), 671–683.

Clopper, C. G., Mitsch, J. F., & Tamati, T. N. (2017). Effects of phonetic reduction and regional dialect on vowel production. *Journal of Phonetics*, *60*, 38–59.

Clopper, C. G., & Pierrehumbert, J. B. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *The Journal of the Acoustical Society of America*, *124*(3), 1682–1688.

Clopper, C. G., & Tamati, T. N. (2014). Effects of local lexical competition and regional dialect on vowel production. *The Journal of the Acoustical Society of America*, *136*(1), 1–4. https://doi.org/10.1121/1.4883478

Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, *30*(1–2), 1–31. https://doi.org/10.1080/23273798.2014.963130

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, *35*(2), 180–209. https://doi.org/10.1016/j.wocn.2006.03.004

Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2010). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, *1*(2), 425–452.

Coupland, N. (1980). Style-shifting in a Cardiff work-setting. *Language in Society*, *9*(1), 1–12.

De Jong, K., Beckman, M. E., & Edwards, J. (1993). The Interplay Between Prosodic Structure and Coarticulation. *Language and Speech*, *36*(2–3), 197–212. https://doi.org/10.1177/002383099303600305

De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America*, *97*(1), 491–504.

Docherty, G. J., Langstrof, C., & Foulkes, P. (2013). Listener evaluation of sociophonetic variability: Probing constraints and capabilities. *Linguistics*, *51*(2), 355–380.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, *41*(1), 87–100.

Eckert, P., & Labov, W. (2017). Phonetics, phonology and social meaning. *Journal of Sociolinguistics*, *21*(4), 467–496.

Erickson, D. (1998). Effects of contrastive emphasis on jaw opening. *Phonetica*, *55*(3), 147–169.

Erickson, D. (2002). Articulation of extreme formant patterns for emphasized vowels. *Phonetica*, *59*(2–3), 134–149.

Ferguson, S. H., & Kewley, -Port Diane. (2007). Talker Differences in Clear and Conversational Speech: Acoustic Characteristics of Vowels. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1241–1255. https://doi.org/10.1044/1092-4388(2007/087)

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America*, *97*(4), 2540–2551. https://doi.org/10.1121/1.411974

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474–496.

Hanson, H. M. (2009). Effects of obstruent consonants on fundamental frequency at vowel onset in English. *The Journal of the Acoustical Society of America*, *125*(1), 425–441. https://doi.org/10.1121/1.3021306

Hay, J., Jannedy, S., & Mendoza-Denton, N. (1999). Oprah and/ay: Lexical frequency, referee design and style. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1389–1392.

Jacewicz, E., Fox, R. A., & Salmons, J. (2006). Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change*, *18*(3), 285–316.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language*, *45*, 229–254.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*(1), 125–156. https://doi.org/10.1515/labphon.2011.004

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3–4), 243–276.

Kristensen, L. B., Wang, L., Petersson, K. M., & Hagoort, P. (2013). The Interface Between Language and Attention: Prosodic Focus Marking Recruits a General Attention Network in Spoken Language Comprehension. *Cerebral Cortex*, *23*(8), 1836–1848. https://doi.org/10.1093/cercor/bhs164

Labov, W. (1981). *Field methods of the project on linguistic change and variation*. Southwest Educational Development Laboratory.

Lai, W., Rácz, P., & Roberts, G. (2020). Experience With a Linguistic Variant Affects the Acquisition of Its Sociolinguistic Meaning: An Alien-Language-Learning Experiment. *Cognitive Science*, *44*(4), e12832. https://doi.org/10.1111/cogs.12832

Lee, E.-K., Cole, J., & Kim, H. (2006). Additive effects of phrase boundary on English accented vowels. *Proceedings of the 3rd Speech Prosody Conference, Dresden, Germany*.

Lewandowski, E. M., & Nygaard, L. C. (2018). Vocal alignment to native and non-native speakers of English. *The Journal of the Acoustical Society of America*, *144*(2), 620–633.

Li, A., & Roberts, G. (2023). Co-Occurrence, Extension, and Social Salience: The Emergence of Indexicality in an Artificial Language. *Cognitive Science*, *47*(5), e13290. https://doi.org/10.1111/cogs.13290

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling* (pp. 403–439). Springer.

McCarthy, C. (2011). The Northern Cities Shift in Chicago. *Journal of English Linguistics*, *39*(2), 166–187.

Mo, Y., Cole, J., & Hasegawa-Johnson, M. (2009). Prosodic effects on vowel production: Evidence from formant structure. *Interspeech 2009*, 2535–2538. https://doi.org/10.21437/Interspeech.2009-668

Munson, B., & Solomon, N. P. (2004). The Effect of Phonological Neighborhood Density on Vowel Articulation. *Journal of Speech, Language, and Hearing Research : JSLHR*, *47*(5), 1048–1058.

Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona1. *Journal of Sociolinguistics*, *11*(4), 478–504. https://doi.org/10.1111/j.1467-9841.2007.00334.x

Pycha, A., & Dahan, D. (2016). Differences in coda voicing trigger changes in gestural timing: A test case from the American English diphthong /aɪ/. *Journal of Phonetics*, *56*, 15–37. https://doi.org/10.1016/j.wocn.2016.01.002

Rickford, J. R., & McNair-Knox, F. (1994). Addressee-and topic-influenced style shift: A quantitative sociolinguistic study. *Sociolinguistic Perspectives on Register*, *235*(276), 10–1093.

Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, *5*, 6–1.

Samara, A., Smith, K., Brown, H., & Wonnacott, E. (2017). Acquiring variation in an artificial language: Children and adults are sensitive to socially conditioned linguistic variation. *Cognitive Psychology*, *94*, 85–114. https://doi.org/10.1016/j.cogpsych.2017.02.004

Sandberg, K. (2024). *The Interpretation of Prosodic Prominence Conveying Contrast and Intensity* [PhD Thesis, NORTHWESTERN UNIVERSITY].

https://linguistics.northwestern.edu/documents/dissertations/2024-03-08_sabdberg_kate_phd_dissertation-submitted.pdf

Schilling, N. (2013). Investigating Stylistic Variation. In *The Handbook of Language Variation and Change* (pp. 325–349). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118335598.ch15

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, *23*(3–4), 193–229. https://doi.org/10.1016/S0271-5309(03)00013-2

Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, *118*(3), 1677–1688.

Sneller, B., & Roberts, G. (2018). Why some behaviors spread while others don't: A laboratory simulation of dialect contact. *Cognition*, *170*, 298–311.

Steffman, J. (2021a). Contextual prominence in vowel perception: Testing listener sensitivity to sonority expansion and hyperarticulation. *JASA Express Letters*, *1*(4), 045203. https://doi.org/10.1121/10.0003984

Steffman, J. (2021b). Prosodic prominence effects in the processing of spectral cues. *Language, Cognition and Neuroscience*, *36*(5), 586–611. https://doi.org/10.1080/23273798.2020.1862259

Sumner, M., Kim, S. K., King, E., & McGowan, K. B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology*, *4*, 1015.

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*(4), 445–472. https://doi.org/10.1016/j.wocn.2006.12.001

Wagner, M. A., Broersma, M., McQueen, J. M., Dhaene, S., & Lemhöfer, K. (2021). Phonetic convergence to non-native speech: Acoustic and perceptual evidence. *Journal of Phonetics*, *88*, 101076.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, *128*(2), 179–186.

Appendix A

Table of written stimuli used in Experiment 1

Table 18: List of stimuli used in Experiment 1

Subject	**Verb**	**Target object**	**Distractor object**
The bee	Goes in	The hive	The acorn
The wheel	Goes on	The bike	The wall
The farmer	Picks up	The seed	The anchor
The goat	Looks like	The sheep	The hammer

The parents	Go with	The kid	The cloud
The waiter	Picks up	The tip	The car
The fish	Goes to	The bait	The key
The teacher	Hands out	The page	The sun
The pillow	Goes on	The bed	The pencil
The dog	Goes to	The vet	The clock
The ball	Goes with	The bat	The oven
The driver	Gets in	The cab	The sandwich
The hiker	Puts on	The boot	The apple
The plate	Goes with	The food	The radio
The frog	Looks like	The toad	The ice cream
The sailor	Gets on	The boat	The pen
The tea	Goes in	The cup	The letter
The soap	Goes in	The tub	The cheese
The flower	Comes from	The bush	The chicken
The shoe	Goes on	The foot	The tree
The pan	Goes with	The pot	The tractor
The pea	Comes from	The pod	The guitar
The people	Live in	The house	The lamp

The milk	Comes from	The cows	The seaweed
The children	Play with	The toys	The couch
The girls	Talk to	The boys	The fork

[i] Both model talkers produced all spoken sentences the same number of times and participants were exposed to all combinations of target phrase and spoken phrase the same number of times. However, a mistake in counterbalancing stimuli resulted in a partial confound between target phrase and model talker in the object focus condition. Rather than both model talkers being associated with all twelve target phrases once, each model talker spoke on a trial associated with four target phrases once, four twice, and four not at all. For example, Talker One's voice was heard on trials where the target phrases was "Word A to word B" twice, "Word A to word C" once, and "Word A to word D" zero times. However, trials were still balanced with regard to how often each model talker produced each word in each position. For example, although listeners may not have heard Talker One speak on trials where the target phrases was "Word A to word D," they still heard Talker One in sentence where the object was "Word D" three times, the same number as for Talker Two. In order for this to affect results, there would need to be a three-way interaction between talker voice, target object, and spoken subject such that participants produced objects differently depending on which speaker produced which specific subject. Because subjects were balanced in phonological forms, and because the frequencies of specific subjects and objects was balanced between talkers, I can not think of any mechanism by which such a interaction would emerge.