

5-2021

Cross-Dialectal Vowel Mapping and Glide Perception

Abram Clear

Follow this and additional works at: <https://scholarworks.wm.edu/honorsthesis>



Part of the [Anthropological Linguistics and Sociolinguistics Commons](#), [Phonetics and Phonology Commons](#), and the [Psycholinguistics and Neurolinguistics Commons](#)

Recommended Citation

Clear, Abram, "Cross-Dialectal Vowel Mapping and Glide Perception" (2021). *Undergraduate Honors Theses*. William & Mary. Paper 1645.

<https://scholarworks.wm.edu/honorsthesis/1645>

This Honors Thesis -- Open Access is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Cross-Dialectal Vowel Mapping and Glide Perception

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Arts in Linguistics from
William & Mary

by

Abram W. Clear

Accepted for Honors
(Honors)

Anya Hogoboom
Professor Anya Hogoboom, Director

Kate Harrigan
Professor Kate Harrigan

Leslie Cochrane
Professor Leslie Cochrane

Adela Amaral
Professor Adela Amaral

Williamsburg, VA
May 10, 2021

Introduction	3
Background	4
2.1. Models of Speech Perception	4
2.2. Dialectal Variation and Speech Perception	5
2.3. Glides	6
2.4. Glide Perception and Perceptual Mapping	7
2.5. Appalachian English	9
2.6. Summation of Background	10
Methods	11
3.1. Participants	11
3.2. Experiment Design	11
3.3. Stimulus Creation	13
3.4. Procedure	15
3.5. Inclusion Criteria	15
Results	16
4.1. Results by Dialect	16
4.2. Statistical Model	17
Discussion	20
5.1. Acoustic Glides	20
5.2. The Existence of Phantom Glides	20
5.3. The [-ja] suffix conditions	22
5.4. Parallel Processes in Speech Perception	23
Future Directions	23
Conclusion	24
Appendix A	24
Appendix B	24
Appendix C	25
References	26

Cross-Dialectal Vowel Mapping and Glide Perception¹

Abram W. Clear

Abstract. Broadening our understandings of how the perceptual system accounts for dialectal vowel variation, this research investigates the perceptual mapping of Appalachian English (AE) monophthongal [aɪ]. I explore this mapping through the secondary perception of palatal glides in hiatus sequences of monophthongal [aɪ.a]. Formant transitions from a high front vowel to a non-high, non-front vowel mimic the formant signature of a canonical [j], resulting in the perception of an *acoustic glide* (Hogoboom 2020). I ask if listeners may still perceive a glide when canonical formant transitions are absent. If participants map monophthongal [aɪ] to a high front position, they might perceive a glide that is not supported by the acoustic signal, which I term a *phantom glide*. Ninety-six participants (45 of which were native AE speakers) heard 30 different English words ending in [i], [ə], or monophthongal [aɪ] (i.e. *tree, coma, pie*) that had been suffixed with either [-a] or [-ja]. They were asked to identify which suffixed form they heard. Participants in both dialect groups sometimes perceived a glide that was truly absent from the speech stream. In these cases, participants mapped static formants in monophthongal [aɪ.a] stimuli to a diphthongal /aɪ/ with a high front endpoint, causing the perception of the necessary F1 fall and subsequent rise of a [j]. Using recent models of speech processing, which encode both social and acoustic representations of speech (e.g. Sumner et al. 2014), I discuss the mapping of monophthongal [aɪ] to a privileged diphthongal underlying form.

1. Introduction

Depending on the heuristic, the United States can be described as home to over 29 dialects (Carver 1987), which non-linguists might think of as accents. Everyone speaks a dialect, even if they insist that they talk “normally” or think of their speech as some unrealistic “standard variety.” While some are stigmatized, no dialect is better, regardless of what prescriptive education, stereotyping, or linguistic prejudice encourages speakers to believe. All dialects have regular rules and patterns and vary by word-choice, grammatical structure, and sound productions, which speakers learn from their parents and home communities during childhood. On the surface, dialects are often defined by the use of

¹ I thank my incredible advisor and friend Anya Hogoboom for her support since this project’s conception. I thank members of the William & Mary Linguistics Labs for their helpful feedback. Leslie Cochrane and Kate Harrigan were invaluable collaborators during experiment design. Appalachian English speakers in my southwestern VA community, eastern TN, and KY are similarly owed my sincerest thanks. Finally, thank you to Kim Love of K. R. Love Quantitative Consulting and Collaboration for statistical consulting and my poster audience at the 2021 meeting of the LSA for their constructive observations.

uniquely salient linguistic features. “Having a dialect” is equivalent to using a specific dialect’s features; dialectal affiliation hinges on production. However, this conception of dialect leaves the perceptual system unexplored.

When processing language, speakers build interpretations out of continuous, chaotic speech streams. These interpretations rely on matching sounds to expected sound and word categories via the language processing system. Dialectal variation, which leads to the introduction of novel, unfamiliar forms in the speech stream, poses a challenge to the processing system. To process dialectal variation, do speakers build new categories into their system or learn to expand an existing category? Does identifying with a dialect affect a speaker’s habitual processing of its varying forms? As language cannot be divorced from context, how do stereotypes and social factors influence the processing of sound variation?

This work is a novel exploration of cross-dialectal vowel perception via the uninvestigated, secondary effect of glide perception. My findings support recent models of speech processing that store social information, like prestige and negative stereotyping, with sound memories that are used for sound perception.

2. Background

2.1. Models of Speech Perception

Countering an established generative approach, exemplar-based models of speech perception hypothesize categories within the perceptual system that are broader than abstracted, lone phonemes. Exemplar-based models assume speech categories are represented in memory by a large cloud of exemplars, which are organized in a cognitive map according to form similarity (Pierrehumbert, 2000). Distributional properties are noted by listeners, and listeners are able to shift boundaries of prelexical categories, representations similar to phonemes in a generative model (Sumner 2011; Cutler et al., 2010). Exemplar-based models can explain realities of speech production that are not compatible with generative, rule-governed approaches. Brief examples of phenomena exemplar models can reasonably account for include ideolectal variation in vowel contrasts (Stollenwerk 1986), differing reduction patterns across homophones (Lavoie 2002), context-improved perception (Picket & Pollack 1963), listener expectations altering perception (Strand 2000), and memory inequalities. In exemplar models, language sound patterns are defined more by extension than by rule (Johnson, 2005).

Exemplar-based models of grammar are popular among sociolinguists and speech perceptualists (e.g. Goldinger 1996, 1998; Johnson, 1997, 2007; Hay et al. 2006; Pierrehumbert 2001, 2006). Exemplar-based models allow for the storing of non-phonetic social factors and nuanced speaker information alongside mental representations; underlying exemplars are fuller and more intricately connected than generative phonemes,

which allows for the integration of more top-down cues in the process of perception. Sumner et al. (2014:2) propose that “socially salient tokens are encoded with greater strength (via increased attention to the stimulus) than both typical and atypical non-salient tokens,” which results in what they term *social-weighting* that explains memory inequality and unequal priming effects. In their expanded exemplar model, acoustic patterns and social representations are mapped in tandem via a dual-route.

This work uses the term “mapping” to refer to a complex set of processes that could be components of either generative or exemplar models of speech perception. “Mapping” may be considered synonymous with “matching” of an exemplar at either the word or sound level. If attempting to fit my findings within a generative model, “mapping” may be equivalent to “abstracting to a prototype” or phoneme. These processes cannot be fully disambiguated as exemplar-based models can, on the surface, “exhibit abstraction behavior” due to exemplar storage (Johnson 2005: 301). I now turn to recent studies of the language systems’ processing of natural variation across dialects and individual speakers. Generally, these studies attempt to disambiguate the presence of one or two underlying representations for the perception of dialectal variants.

2.2. *Dialectal Variation and Speech Perception*

Studying the processing of r-ful (ie. bak[ə]) and r-less (ie. bak[ə]) surface forms in populations of NYC English and General American (GA)² English, Sumner and Samuel (2009) argued for the presence of different phonological representations for different dialectal forms. Overt³ speakers of New York City (NYC) English produced r-less forms but used both r-less and r-ful forms in perception, indicating that they store representations for both variants. In an experiment using long-term repetition priming, either form was able to facilitate the recognition of the other. Covert⁴ speakers of NYC English and GA speakers only encoded the r-ful forms in the long-term priming task; they saw no facilitative effects with r-less forms. While r-less forms were never effective primes for GA speakers, Covert and Overt NYC speakers were facilitated by r-ful and r-less forms in other form and semantic priming tasks. The authors made a distinction between bidialectal speakers (Overt NYC speakers) and *fluent listeners* (Covert NYC speakers). Bidialectal speakers map NYC r-less and GA r-ful surface variants to separate lexical representations; fluent listeners easily encode surface variants as realizations of a single, r-ful underlying form. Beyond evidence for dialect specific encodings, Sumner and Samuel found a general benefit for unmarked variants (r-ful forms in their study), which supported previous findings allocating more perceptual weight and facilitative capacity to prestigious forms (McLennan et al. 2003; Ranbom & Connine 2007). In a later study, Sumner and Kataoka

² These researchers use the term “General American” to refer to speakers with r-ful vernacular forms who do not “exhibit any other regionally marked characteristics” (488).

³ r-less producing speakers who were born and raised in the r-less NYC dialect region

⁴ r-ful producing speakers who were born and raised in the r-less NYC dialect region

(2013) found a memory benefit for the same r-less form produced by a speaker of a prestigious dialect, non-rhotic British English.

Investigating the perception of variation in the language system of Basque-Spanish bilinguals, Samuel and Larraza (2015) argued for a dual-mapping route to a single exemplar, resulting in a quasi-allophonic representation. Basque contrasts the affricates /tʃ/ and /tʃ/. Spanish only has /tʃ/, and L1 Spanish speakers often report difficulty with the /tʃ/ form. Teaching Basque-Spanish bilinguals novel words with a consistent form of either /tʃ/ or /tʃ/, they controlled the encoding of single representations into the lexical system of their participants. No participants encountered /tʃ-/tʃ/ variation for any novel word. Samuel and Larraza tested the speakers' acceptability judgements of "mispronounced" variants, substituting /tʃ/ for /tʃ/ and vice versa. The "mispronounced" variants were accepted as real Basque words with an error rate of ~30%. Another experiment revealed consistent discrimination of these sounds by participants outside of a lexical context. For purposes of lexical access, [tʃ] and [tʃ] are treated as essentially allophonic variants for Basque-Spanish bilinguals. This underspecified model of dual-mapping to a single exemplar mechanically parallels the proposed language system of a *fluent listener* (Sumner & Samuel 2009).

Studying American English vowel perception, Fridland and Kendall (2011) concluded that "processing is affected by both what you say and what others around you say" (792). They presented participants from the U.S. South, North, and West with a forced choice vowel identification task. Participants were presented with a word containing a synthesized vowel at one of seven steps between a sample speaker's natural endpoints of /e/ and /ɛ/. Participants were asked to identify the word as its /e/ or /ɛ/ form (ie. *bait* or *bet*). The authors also collected production data from a subset of their participants via a reading passage and word list. Most important for this research, Southern speakers' productions were influenced by the Southern Vowel Shift; they produced the mid-front vowels, /e/ and /ɛ/ much closer together than their Northern and Western counterparts. Reflecting a perceptual system influenced by their own habits of production, Southerners were found to maintain /e/ perception along the continuum of synthesized vowels toward the /ɛ/ target longer than non-Southerners.

2.3. *Glides*

Glides are semi-vowel consonants whose articulation is very similar to movement common in vowel-to-vowel transitions. They are described as "maximally similar to their vocalic environment" (Uffmann 2007:458). Past attempts to differentiate phonological glides from vowels in succession cite degrees of constriction, noting consonantal glides typically possess more frication and constriction than V.V sequences (Padgett 2008). Recent research has called this heuristic into question. For example, Burgdorf and Tilson (2021)

found high vowels to be significantly more constricted than glides. Using electromagnetic articulography, they also found glide productions to be more precise, having less variance in articulation than vowel productions. Regardless of fine-grained acoustic properties, glides and certain vowel sequences are perceptually confusable.

As the acoustic realization of glides mimics vowel transition, certain vowels in hiatus position (i.e. adjacent but parsed into separate syllables) can precipitate the acoustic waveform of and perception of a glide. Formant transitions from a high front vowel to a central or low back vowel mimic the formant signature of a canonical [j]. Though one lacks an intentional glide, forms like [ki.ə] and [ki.jə] are confusable to English speakers. Following Hogoboom (2020), I call these unintentional glides incurred from vowel transition *acoustic glides*.

This line of research partially investigates if I can perceptually disambiguate acoustic glides from phonologically present glides. Cross-linguistically, glide formation, /CV₁.V₂/ → [CGV₂], is a commonly employed method of hiatus resolution (Casali 2011). Homorganic glide epenthesis/formation has been argued to be less phonologically costly than the epenthesis of a new consonant segment (Uffmann 2007). Hogoboom (2020) argues phonological glides and high front to central or low back vowel hiatus sequences are confusable even for speakers of languages that contrast [i.a] and [i.ja] forms (i.e. Mandarin, Korean, etc.).

2.4. *Glide Perception and Perceptual Mapping*

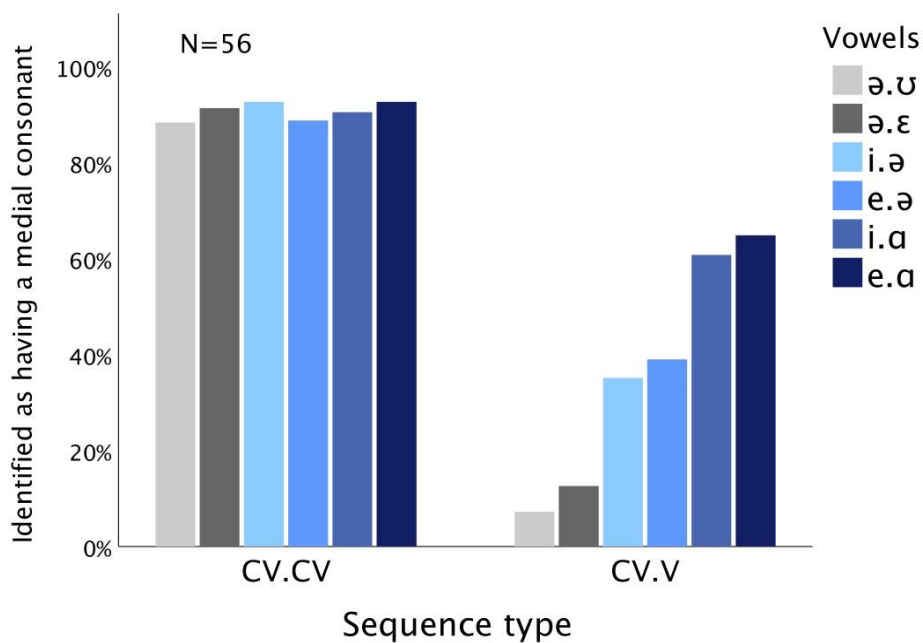
Hogoboom (2020) established the confusability of acoustic glides and phonological glides. The rise of F1 and fall of F2 necessary to transition from [i] to [a], [o], or [ə] can incur the perception of an acoustic glide. I ask, could the *perception* of a high front starting position in a V.V sequence incur the reporting of a glide that is fully absent from the speech signal. Rather than present through formant mimicry or acoustic byproducts, I investigate the perception of glides that lack any acoustic foundation. I call these consonants that may be perceived via perceptual mapping and top-down processes of lexical access *phantom glides*. I hypothesize that phantom glides may be perceived by accessing the underlying representation of a high front vowel (be that a sound/word-level exemplar or phoneme) when presented with a non-high vowel's surface form.

This hypothesis is supported by previous research (Clear & Hogoboom, in prep). To investigate the likelihood of acoustic glides being perceived and identified as real consonants, we crafted a forced-choice perception task with German consonants and German vowels (MBROLA, Dutoit et al. 2019). Participants heard two-syllable CV.CV and CV.V tokens and were asked “Was there a consonant between the vowels?”. Consonants unused in English were synthesized vowel-medially in CV.CV cases. In addition to discerning the confusability of acoustic glides with an intentional, phonological

consonant, the experiment was designed to further our understanding of which vowel combinations reliably give rise to the perception of acoustic glides.

Results are shown in Figure (1). Sequences with intervening German consonants (left aligned bars in (1)) were reliably perceived as such. Vowel sequences without a synthesized intervening consonant (right aligned bars in (1)) varied in reported consonant perception. Lacking any glide-like formant transition, [ə.ʊ] and [ə.ɛ] cases were rarely perceived with intervening consonants. As expected, [i]-initial vowel sequences incurring unintentional glide-like formant transition led to the perception of acoustic glides. Unexpectedly, participants were as statistically likely to report hearing an intervening consonant in the cases that started with [e], a static mid front vowel, as in those that started with [i]. The German, monophthongal [e] used in stimuli sequences did not have a high front endpoint, which precipitated perception of acoustic glides via subsequent transitions. Though unattested in the speech stream, participants perceived glides in mid front to central and low back sequences at rates comparable to acoustic glide stimuli.

Figure 1. Results indicating phantom glide perception (Clear & Hogoboom, in prep)



We posit that the perception of phantom glides in German [e.ɑ] and [e.ə] stimuli was due to perceptual mapping of the [e] to a higher and fronter form. American English speakers produce [e] as an inherent diphthong [eɪ]. They hear [eɪ] far more frequently than [e] and have no social impetus to privilege an /e/ form. Mapping the monophthongal German [e] to their inherently diphthongal underlying /eɪ/, English speakers perceived a higher and fronter [e] endpoint than was present in the acoustic signal of [e.ɑ] or [e.ə]

stimuli. Perception of a high front endpoint allowed for subsequent perception of formant transitions that mimic a real [j]. This mapping and “imagined distance” gave rise to phantom glides in sequences beginning with a mid front vowel.

This research pushes the question of perceptual mapping further. I ask if speakers can map an initial low front vowel in a hiatus sequence to a high front underlying form. This high front mapping could lead to the perception of formant transition in relatively static low vowel to low vowel hiatus sequences, which could incur the secondary effect of phantom glide perception. I use the monophthongal form of /aɪ/, which is discussed in detail below, to ascertain the existence of phantom glides. Due to the possible encoding of /aɪ/ as a prestige form over monophthongal /aɪ/ or the quasi-allophonic positioning of monophthongal [aɪ] in perceptual relation to diphthongal /aɪ/, listeners have the potential to map monophthongal [aɪ] to a diphthongal underlying form. Mapping to a diphthongal underlying form (with a high front end position) could incur the perception of phantom glides in certain hiatus contexts.

2.5. *Appalachian English*

Appalachian English (AE) is a dialect spoken throughout the southern portion of the Appalachian mountain range, including western North Carolina and Virginia, southern West Virginia, and Eastern Kentucky and Tennessee (Wolfram 1977). Common linguistic features of the dialect include: a-prefixing (i.e. *He was a-workin'*), double modals (i.e. *I might could do that*), was-leveling (i.e. *They was goin' over there*), nonstandard past tense (i.e. *I knowed that already*), expletive they (i.e. *They's copperheads around here*), nonstandard intensifiers (i.e. *right, awful, plumb*), demonstrative them (i.e. *Hand me them flowers*), positive anymore (i.e. *She's more northern than she is southern anymore*), multiple negation (i.e. *ain't nothing*), and unique lexical items (i.e. *reckon, liketa, holler, poke*, etc.) (Wolfram & Christian 1976). Known for their advanced stages of the Southern Vowel Shift, Appalachian English speakers produce front tense and lax vowels in overlapping or canonically reversed target positions (Irons 2007). The vowels /eɪ/ and /ɛ/ appear to have reversed positions for many speakers, and Irons (2007) found that a similar /ɪ/ and /i/ reversal is progressing in real time. It is important to recognize this speaker population is not monolithic; some research has supported the differentiation of Appalachian English into separate *Appalachian Englishes* to better capture a multiplicity of linguistic and social identities within the Appalachian mountain range (Hazen 2020).

Most important to this research, Appalachian English speakers produce /aɪ/ as a monophthong with a low front target region (Reed 2014). Monophthongization of the /aɪ/ diphthong has been observed in speakers of broadly Southern American English (e.g., Bernstein 2006; Labov, Ash, & Boberg 2005; Greene 2010). However, it is far more

common in southern Appalachian English productions, appearing in pre-voiceless consonant contexts as well as in pre-voiced and word-final environments (Thomas 2001).

Studies in perceptual dialectology (folk linguistics), have shown Southern speech features to be the most regionally marked and salient to non-linguists in the United States (Clopper & Pisoni 2004, 2007; Niedzielski & Preston 2003; Preston 1989, 1993). As Appalachia is stigmatized, most Appalachians have a high degree of linguistic awareness; they are conscious of how their language is negatively perceived by non-Appalachians (Greene 2010; Reed 2014, 2016). While possessing covert prestige within the region, Appalachian English features can index negative stereotypes of unintelligence, closed-mindedness, and depravity. Best summarized by Williams (2002), the negative stereotype faced by Appalachians is one of “the profligate hillbilly, amusing but often also threatening, defined by a deviance and aberration, a victim of cultural and economic deprivation attributable to mountain geography” (17).

Such prejudice encourages assimilation to more “standard,” prestigious linguistic forms, which may be produced by covert, diasporic Appalachians through dialectal code-switching. Reed (2018) has investigated the effects of *rootedness*, which he defines as one’s attachment to place, on monophthongal [aɪ] production. Reed (2018) found nearly categorical change in one speaker’s monophthongal [aɪ] production after she left Appalachia. The diphthongal, mainstream production of [aɪ] was “more fitting for her current social identity” and likely helped evade negative connotations of Appalachian identity (Reed 2018:421). Conversely, some native Appalachians retain Appalachian English speech forms in new social contexts for a sense of connection to their linguistic and cultural roots (Hazen & Hamilton 2008).

2.6. *Summation of Background*

A synthesis of perceptual and dialectal variance, this study hopes to push the bounds of perceptual mapping’s known effects. Can perceptual mapping lead to the secondary perception of glides which are fully absent from the sound signal in vowel hiatus sequences? Does this perception differ based on speaker dialect? If phantom glides are perceptible, I expect Appalachian English speakers to perceive them less frequently, as they presumably privilege a vowel mapping to the static, monophthongal form with which they identify. Alternatively, competency with the prestigious, diphthongal form (and its establishment as a perceptually salient, ideal) may affect their preferred mapping (Sumner et al. 2014) and therefore cause the secondary effect of phantom glide perception at an equal level to non-Appalachian speakers. The stigmatization of Appalachian English forms and salient contrast to academic, prestigious, hyper-“correct” forms may affect Appalachian English speakers’ mappings of their own monophthongal vowels.

3. Methods

To determine the limits and robusticity of perceptual mapping's effects, this study placed Appalachian English's static, low front monophthongal [aɪ]⁵ in hiatus position with [a]. Participants perceived glides in monophthongal [aɪ.a] sequences only if their perception was influenced by perceptual mapping to a diphthongal form, which incurred the perceived formant transitions necessary to precipitate a phantom glide. To better determine how perceptual mapping might be influenced by a speaker's native dialect, this study compares Appalachian English speakers with speakers of other American English dialects.

3.1. *Participants*

Ninety-six native English speakers living within the continental United States participated in the study. Identifying themselves as native Appalachians, 45 participants acquired their first language within the geographic confines of Appalachia and were coded as Appalachian English speakers (mean age = 41, age range = 21-74, 6 cisgender men, 35 cisgender women, 3 of unspecified gender identity). 51 participants did not categorize themselves as native Appalachians and were coded as speakers of other United States dialects (mean age = 30, age range = 19-73, 11 cisgender men, 32 cisgender women, 1 trans man, 2 nonbinary, 1 agender, 4 of unspecified gender identity). Initial participants were recruited through social media advertising (Facebook) and the University of Kentucky's Appalachian Studies listserv. Subsequent Appalachian English participants were recruited through snowball sampling, which entailed asking previous participants to recommend the study to their friends and family members.

3.2. *Experiment Design*

Designed as a forced-choice perception task, participants were played a series of stimuli and asked to identify each as one of two options. The stimulus was always an English word suffixed with either [-a] or [-ja]; participants were asked to indicate which suffixed form of each word they heard. As the root words were all vowel-final, the [-a] suffix condition created a hiatus configuration, creating the environment for the perception of acoustic or phantom glides.

The stimuli consisted of three types of words: ten words ending in each of [i] (type 1), [ə] (type 2), and monophthongal [aɪ] (type 3) were included, giving a total of 30 words. See Appendix A for a complete list of stimulus words by type. The stimulus types were designed to give rise to the perception of an acoustic glide (type 1), no glide (type 2), or a potential phantom glide (type 3). Stimulus types were designed to create optimal acoustic

⁵ Although [aɪ] is not an accurate transcription of monophthongal [aɪ], I will continue to refer to the vowel realization as "monophthongal [aɪ]." I make this choice to show its connection to the diphthongal form and avoid confusion with [a], or [æ], which might be more phonetically accurate transcriptions.

and phantom glide environments specific to the [-a] condition. The [-ja] suffix condition presented participants with an intentional, phonological glide consonant following word-final vowels. See Figure (2) for a breakdown of comparable populations, stimulus types, and conditions.

Figure 2. Experiment conditions breakdown

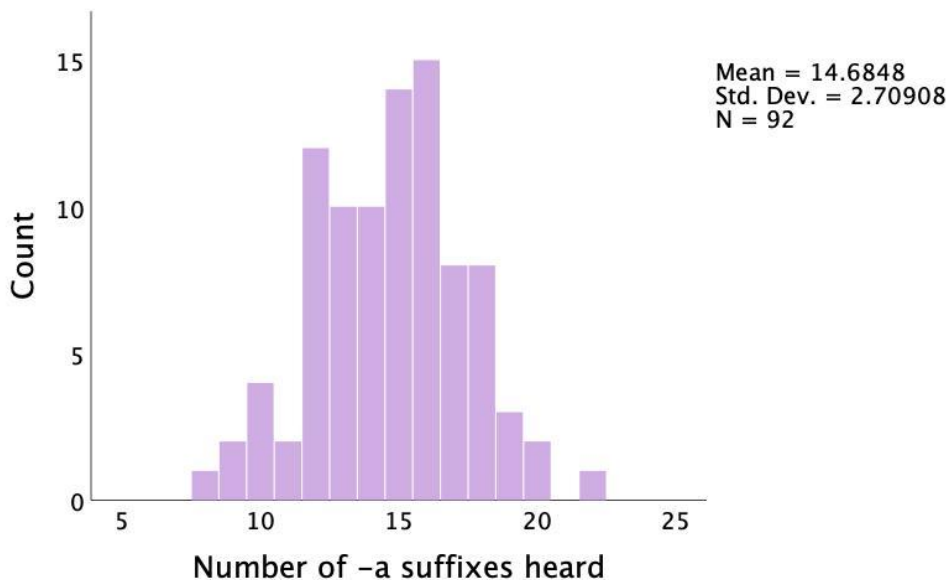
Comparable Populations	Stimulus Types	Conditions
Appalachian English Speakers	① i#	-a
		-ja
	② ə#	-a
		-ja
	③ mono. aɪ#	-a
		-ja
Other American English Dialect Speakers	① i#	-a
		-ja
	② ə#	-a
		-ja
	③ mono. aɪ#	-a
		-ja

Type 1 stimuli (i.e. *bee, flea, knee*) were included to measure the perception of acoustic glides. In the [-a] suffix condition, type 1 stimuli were likely to incur the perception of acoustic glides due to glide-like formant transitions from [i] to the suffixed [a]. Formant transitions from the high, front [i] to the low, back suffixed [a] mimic a phonological glide in the acoustic signal. Type 2 stimuli (i.e. *sofa, tuna, villa*) were included as controls. When confronted with type 2 stimuli, participants were expected to perceive few glides. The formant transitions from the mid, static [ə] to the low, back suffixed [a] are not drastic enough, and they do not start from a high enough initial articulation to acoustically resemble a glide. Nor did participants have an impetus to map [ə] to a high, front position to incur phantom glide perception. Most directly informative for my research question, type 3 stimuli (i.e. monophthongal *pie, spy, lie*) were included to test an extreme case of perceptual mapping. A truly low vowel to low vowel sequence, type 3 monophthongal [aɪ] to suffixed [a] stimuli lack the notable F1 formant transition necessary for the perception of an acoustic glide. Therefore, phantom glides perceived from type 3 stimuli suffixed with [a] were the result of perceptual mapping. To perceive a glide, participants mapped a low, front, monophthongal [aɪ] to a high, front position, resembling its diphthongal pronunciation.

The [i] and monophthongal [aɪ]-final stimuli were one syllable English words to avoid stress confounds. The control, type 2, could not be one syllable. Due to English's minimum word requirements, English [ə]-final words must be at least two syllables long. Type 2 stimuli therefore contained an unstressed syllable in between their main stress and the suffix.

A script randomized which condition of each stimulus, either the [-a] or [-ja] suffixed form, was played for each participant. Thus, some participants heard more of one condition. As participants were asked to identify 30 suffixed words, one would expect a general split between 15 [-a] and 15 [-ja] forms after randomization for each participant. Though clustered around 15 [-a] suffixed forms for each participant, Figure (3) shows a broader distribution.

Figure 3. Histogram of [-a] suffixes encountered by participants

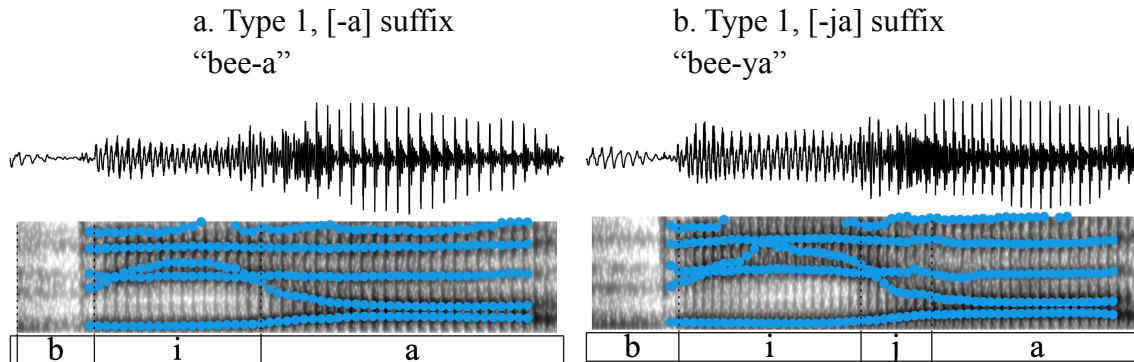


3.3. Stimulus Creation

In total, 60 stimuli were created for this study. All stimuli were recorded by the author (a 21-year-old native Appalachian English speaker). While producing other stimulus sequences using an approximation of standard American English (*SAE), I actively codeswitched to my native Appalachian English to produce stimuli containing monophthongal [aɪ]. For naturalistic formant transitions, each of the 30 English test words was recorded with individual realizations of the [-a] and [-ja] suffixes. The stimuli were recorded in question-answer sentence frames to allow for standard production without prosodic variation (e.g. Which [stimulus] did your sister see? My sister saw the [stimulus] that needed painting). Following recording, stimuli sequences were isolated in Praat

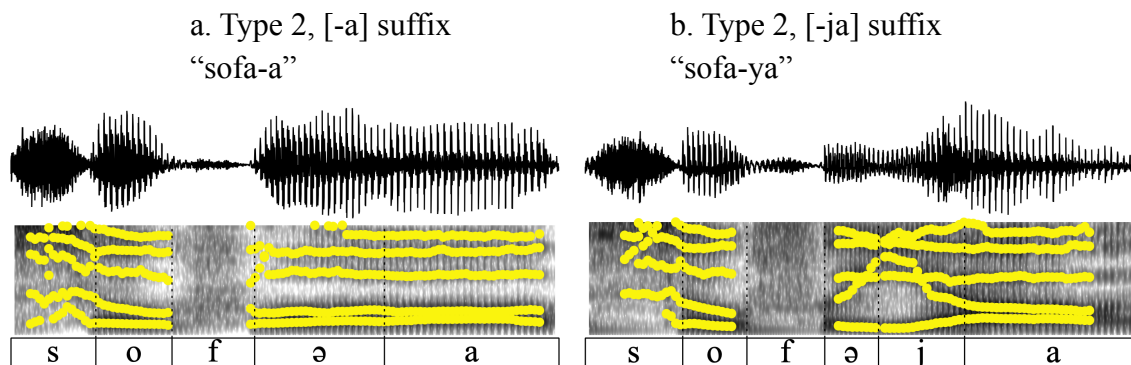
(Boersma & Weenink 2019). Spectrograms of examples of each stimulus type in each condition are shown below. Lines delineating sounds are approximate.

Figure 4.



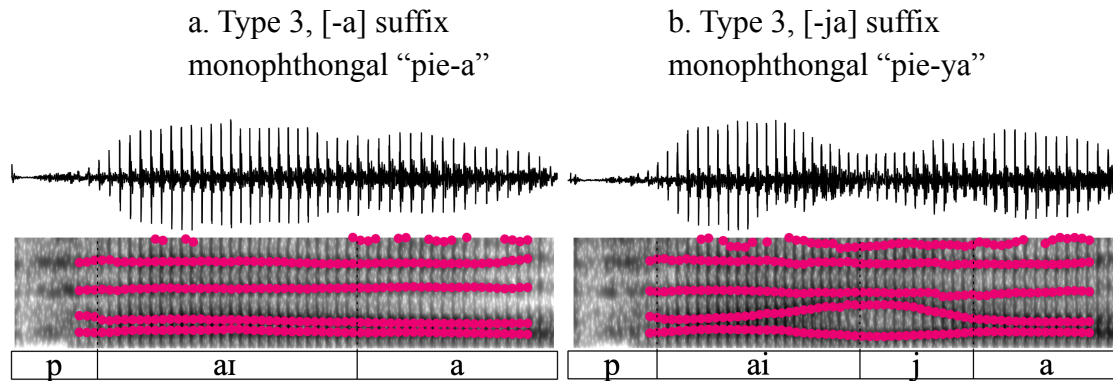
The spectrograms in Figure (4) exemplify type 1 stimuli. While their formant signatures mimic each other, “bee-ya”, shown in Figure (4b), was produced with an intentional glide and “bee-a”, shown in Figure (4a) was produced as a two-vowel sequence. The rise of F1 and fall of F2 necessary to transition from [i] to [a] in Figure (4a) incurs an unintentional acoustic glide. The inherent similarities between the spectrograms reflects the forms’ reasonable confusability.

Figure 5.



Examples of type 2, control stimulus Figures (5a) and (5b) are distinctly different. Figure (5a), a type 2 stimuli suffixed with [-a], shows relatively little formant transition between the word-final [ə] in “sofa” and the suffixed [a]. Nothing in the formant signature remotely mimics a glide. Figure (5b) “sofa-ya”, a type 2 stimulus suffixed with [-ja], is notably different. Following [ə], F1 falls and F2 dramatically plateaus at a high value (indicating a fronter tongue position). F2 proceeds to fall and F1 rises throughout the [j]’s production. As the [j] in Figure (5b) is phonological and intended, these formant movements are expected.

Figure 6.



Finally, Figures (6a) and (6b) show spectrograms of type 3 stimuli. A monophthongal pronunciation of “pie” suffixed with [-a] is shown in Figure (6a). Note the static formants throughout the monophthongal [ai] production that remain nearly unchanged for a transition to [a]. Nothing in the acoustic signal indicates the presence of a glide. Figure (6b) shows a monophthongal [ai] suffixed with [-ja]. While the [ai] formants are initially stationary, an unavoidable F1 fall and F2 rise surface as the speaker transitions to the high, front starting position of a phonological glide. The [j] in Figure (6b) exhibits the archetypical rising F1 and falling F2 of a palatal glide.

3.4. Procedure

This forced-choice task was deployed online through Ibex Farm (Drummond 2016). All participants took part in the study via a weblink without proctoring. They were asked to wear headphones. Following a declaration of consent, participants had the opportunity to play and replay a test sound, which allowed for volume adjustment. Participants were then instructed to assume for the purposes of the study that English has two new suffixes, “-a” and “-ya.” They were told they would be asked to indicate which suffix had been added to 30 different English words. Though encouraged to listen closely, they were asked to go with their first instinct when deciding. Upon hearing each stimulus, two clickable words appeared on the participants’ screens. One click-able word was suffixed with [-a] (e.g. “tree-a”) and the other was suffixed with [-ja] in typical English orthography (e.g. “tree-ya”). Stimuli were not replayable. After making a choice, the screen immediately advanced to the next stimulus, which played without further prompting. The question “Which word did you hear?” remained on the screen for the duration of the study. After identifying their 30th word, participants were presented with a series of demographic questions (see Appendix B) before exiting the study.

3.5. Inclusion Criteria

Only participants who correctly identified 80% of the suffixes when presented with the control type (type 2) were included in this analysis. Two native Appalachian English

speakers were omitted, leaving 43 for analysis. Two speakers of other U.S. dialects were omitted, leaving 49 for analysis.

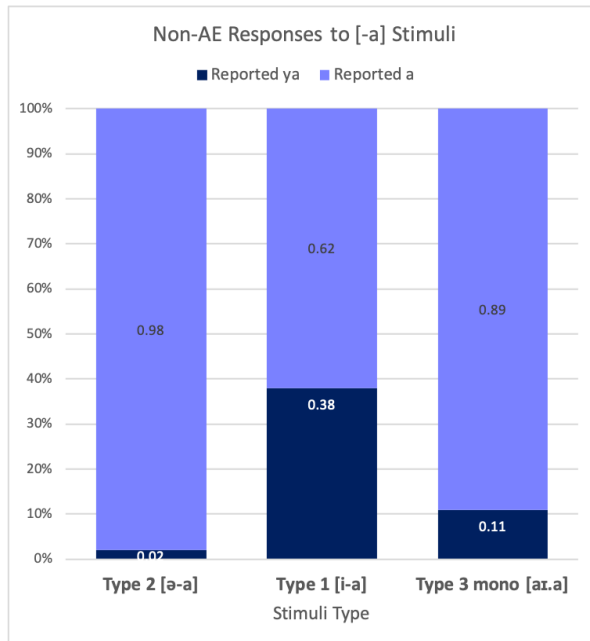
4. Results

4.1. Results by Dialect

The following figures show responses by participant dialect. The graphs are limited to the [-a] suffix condition, which, in potentially causing the perception of acoustic glides and phantom glides, was my most important stimulus configuration. Responses to [-ya] stimuli can be seen in (4.2) and are discussed in (5.3).

Figure 7.

a. Non-AE Responses to [-a] Stimuli (n=49)



b. AE Responses to [-a] Stimuli (n=43)

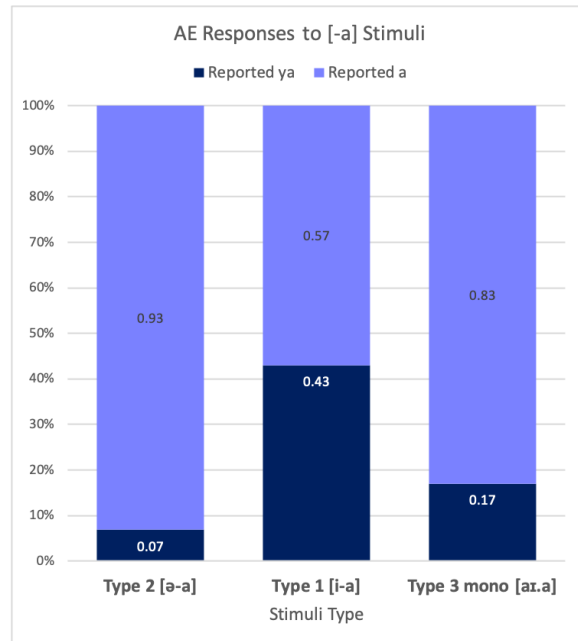


Figure (7a) shows the responses of Non-Appalachian English participants to [-a] stimuli. Participants reported hearing glides in 2% of the type 2 (control) sequences. As expected, Non-Appalachian English participants reported hearing the most glides in type 1 (acoustic glide) sequences. Glides were perceived 38% of the time in those [i-a] sequences. Participants perceived glides in type 3 (phantom glide) sequences 11% of the time. Of particular interest is the fact that notably more glides were perceived in type 3 (phantom glide) sequences than in type 2 (control) sequences. Incurring the perception of the formant transition necessary to perceive phantom glides, Non-Appalachian English participants

sometimes mapped monophthongal [aɪ] as having a high front end point like diphthongal /aɪ/, and thus did in fact hear phantom glides.

Figure (7b) shows the responses of Appalachian English participants to [-a] stimuli. Participants reported hearing glides in 7% of the type 2 (control) sequences. As expected, participants also reported hearing the most glides in type 1 (acoustic glide) sequences. Glides were perceived 43% of the time. Participants perceived glides in type 3 (phantom glide) sequences 17% of the time. Once again, notably more glides were perceived in type 3 (phantom glide) sequences than in type 2 (control) sequences. Thus, Appalachian English participants also must have mapped monophthongal [aɪ] as having a high front end point like diphthongal /aɪ/. This incurred the perception of the vowel distance and formant transitions necessary to perceive phantom glides.

4.2. Statistical Model

A generalized linear mixed effects model (GLMM) approach to logistic regression was run in SPSS with *response* as the dependent variable and *participant dialect*, *stimuli type*, and *suffix condition* as fixed effects. The model included random intercepts for the individual participants and for the individual words. Figure (8) details which terms in the model were statistically significant.

Figure 8. Fixed Effects table

Source	F	df1	df2	Sig.
Corrected Model	42.083	11	342	.000
Dialect	1.389	1	151	.240
Type	18.305	2	34	.000
Suffix	413.434	1	2,748	.000
Dialect*Type	1.362	2	2,748	.256
Dialect*Suffix	5.425	1	2,748	.020
Type*Suffix	139.878	2	2,748	.000
Dialect*Type*Suffix	1.048	2	2,748	.351

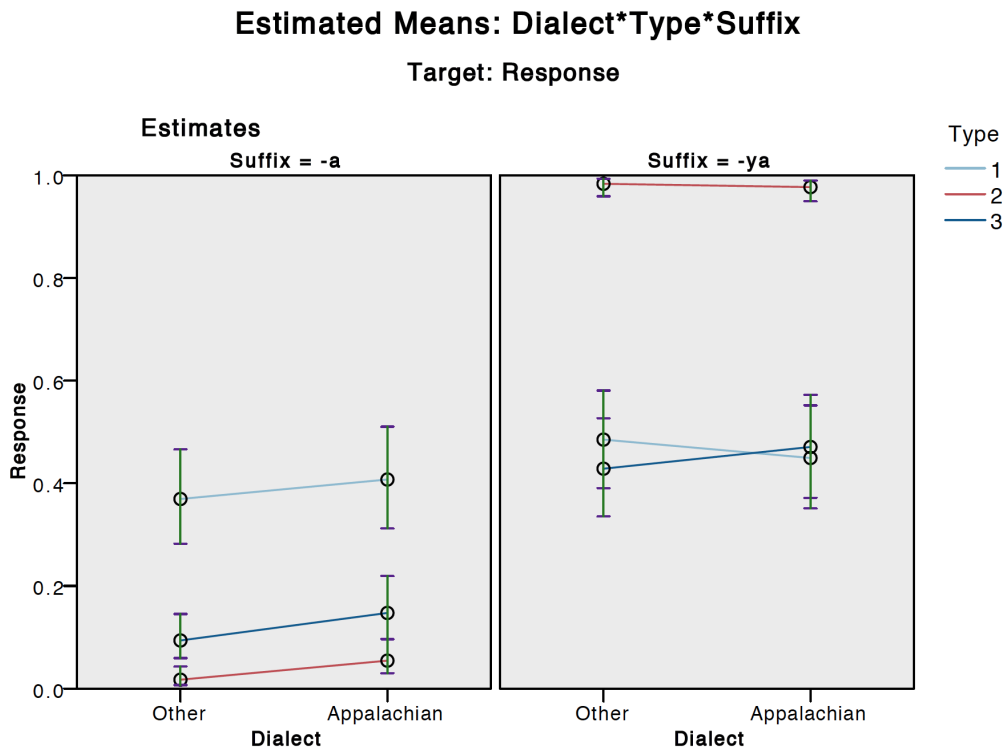
Probability distribution:Binomial
Link function:Logit

As seen above, the three-way interaction between *participant dialect*, *stimuli type*, and *suffix condition* was not significant. The two-way interaction between dialect and type was also not significant. Together, these results do not support the initial hypothesis that participants respond to my type 3 stimuli differently based on their dialectal affiliation. I conclude there was no difference in the perception of any stimulus type due to participant dialect. See Appendix C for the raw, cross-dialectal response data.

The interaction of *dialect* and *suffix* was statistically significant ($p = .020$). Overall, non-Appalachian English speakers were somewhat more likely than Appalachian English speakers to report hearing “-a” in the [-a] suffix condition and “-ya” in the [-ja] suffix condition. This result does not relate to any aforementioned hypotheses and will not be commented on further.

Supporting basic hypotheses that influenced the experiment’s design, the interaction of type and suffix proved statistically significant ($p < .001$). The responses to different suffixes varied greatly by type. Figure (9) graphs the estimated means for the three-way combination of *participant dialect*, *stimulus type*, and *suffix condition*.

Figure 9. Graph of Estimated Means (three-way interaction)



The estimated probability of participants reporting “-ya” for each type and suffix combination is shown on the y-axis. The graph at left shows probability estimates for the

[-a] suffix condition, and the graph at right shows the probability estimates for the [-ja] suffix condition. Stimulus types are differentiated by line color.

Across both graphs, the type 2 (control) condition behaves as expected. Referring to the red line, note the low probability of participants responding “-ya” when played a type 2 stimulus suffixed with [-a] (at left). Note the exact, expected reversal of type 2 responses in the graph at right. Participants are expected to reliably respond “-ya” when played type 2 stimuli suffixed with [-ja]. As the responses to [-ja] suffixes are not informative of the perception of acoustic or phantom glides, I will now focus exclusively on the “-a” suffix estimated means (graph at left). Estimated type 1 stimulus responses, shown with a light blue line, appear to indicate confusion attributed to the perception of acoustic glides. If played a type 1 stimulus with an [-a] suffix, the model predicts participants will categorize it as “-ya” around 40% of the time. The type 3 estimated responses, which are indicated in dark blue, show less expected “-ya” responses than type 1 stimuli but notably more “-ya” responses than type 2 stimuli. The dark blue line is statistically significantly higher ($p < 0.001$; pairwise comparison) than the red line (which refers to responses from the type 2, control condition) showing that phantom glides were in fact perceived. The lack of significant three-way interaction term shows us that phantom glide perception occurred across participant dialect groups.

Pairwise contrasts comparing responses of each stimulus type within each suffix condition are shown in Figure (10).

Figure 10. Pairwise Contrasts

Suffix	Type Pairwise Contrasts	t	df	Significance
-a	1 - 2	± 9.513	185	0.000
	1 - 3	± 7.081	46	0.000
	3 - 2	± 4.365	221	0.000
-ya	1 - 2	± 11.806	264	0.000
	1 - 3	± 0.370	27	0.715
	2 - 1	± 11.990	266	0.000

Responses to the [-a] suffix are significantly different across all three stimuli types. The stimuli types, which incur the perception of nearly no glides (type 2, control), acoustic glides (type 1), and phantom glides (type 3), have statistically significantly different

responses. Perhaps surprisingly, there is no contrast between type 1 and type 3 stimuli within the [-ja] suffix condition. This deviation will be discussed in section (5.3).

5. Discussion

5.1. Acoustic Glides

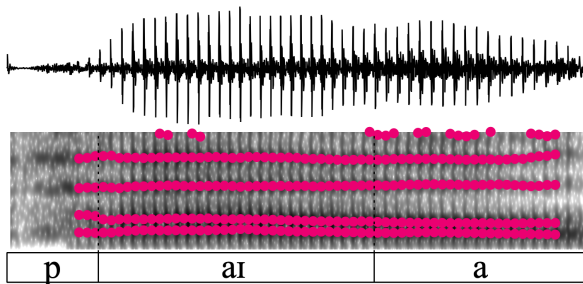
As expected by Hogoboom’s (2020) findings, acoustic glides were confusable with phonological glides for all participants. Across dialects, type 2 (acoustic glide) stimuli suffixed with [-a] were perceived as suffixed with [-ja] in 40% of instances. Similarly, type 2 stimuli suffixed with a phonological glide, [-ja], were perceived as suffixed with [-a] in 53% of instances. Perception of a glide in an [i.a] sequence appeared close to chance; participants were largely unable to perceive a difference between type 2 phonological glide ([-ja] suffixed) and acoustic glide ([-a] suffixed) stimuli. The rise of F1 and fall of F2 necessary to transition from [i] to [a] in hiatus sequences too closely mimics the rise of F1 and fall of F2 in the articulation of a phonological [j] for listeners to differentiate the intended form.

5.2. The Existence of Phantom Glides

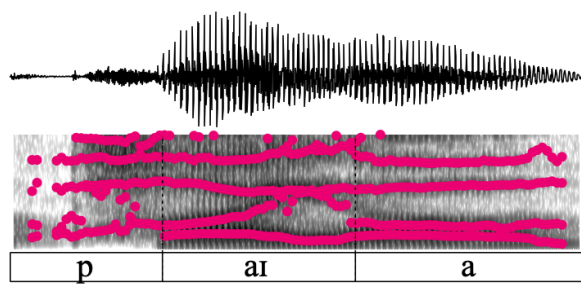
Averaged across all participants, type 3 (phantom glide) stimuli suffixed with [-a] were perceived as suffixed with [-ja] in 14% of instances. Participants in both dialect groups perceived a glide that was truly absent from the speech stream. To perceive phantom glides, participants from both dialects at times mapped static formants in monophthongal [aɪ.a] stimuli (11a) to a diphthongal [aɪ] with a high front endpoint (11b), which incurred the necessary F1 fall and rise for glide perception.

Figure 11.

a. Spectrogram of monophthongal [aɪ.a] stimuli



b. Possible perceptual mapping of monophthongal [aɪ.a] stimuli



Given that speakers can map monophthongal [aɪ] to diphthongal /aɪ/, we might wonder why this does not happen more frequently. The lack of more robust phantom glide perception may support the conception of a dual-representation model, wherein both Appalachian English and Non-Appalachian English speakers have underlying forms for monophthongal and diphthongal /aɪ/. Using exemplar-based logic, Non-Appalachian speakers have access to monophthongal /aɪ/ exemplars through media, cross-dialectal contact, and, as stated, may subconsciously use /aɪ/ production as a heuristic for identifying Southern American English speakers (Clopper & Pisoni 2004, 2007). Similarly, though of a distinct power differential, Appalachian English speakers have access to diphthongal /aɪ/ exemplars through educational settings, mass media, and cross-dialectal contact.

We therefore assume both populations have access to underlying representations of monophthongal and diphthongal /aɪ/. This means both populations can correctly map the monophthongal [aɪ]. However, participants sometimes mapped it to an erroneous high front position, incurring the perception of phantom glides. This mapping outweighed phonetic detail from a continuous speech stream. I propose the diphthongal underlying representation has a prestige pull in the perceptual system through social-weighting, consistent with Sumner et al.'s (2014) proposal. Their model uses social-weighting to explain memory inequality and unequal priming effects. Acoustic patterns and social representations are mapped in tandem via a dual-route in their expanded exemplar model. When presented with a word form or utterance, listeners retain both fine-grained acoustic properties of the form and social evaluations of the form's speaker through a dual-route of *socioacoustic-encoding*. Though potentially atypical or of lower vernacular frequency, socially salient or idealized forms (like those of a perceived standard) are strongly encoded in the perceptual system. These stronger encodings are more easily accessed during speech processing.

I posit social-weighting of diphthongal /aɪ/ lures participants to make erroneous mappings to its percept when processing speech, as opposed to mapping to monophthongal /aɪ/. Diphthongal /aɪ/ percepts retain higher degrees of saliency in the lexical access and speech processing system due to the qualitative social information stored about the form. Consistently produced in careful speech and in the speech of speakers using a perceived Standard American English,⁶ diphthongal /aɪ/ is encoded as an idealized form. This qualitative idealization leads to its weight over a monophthongal counterpart in the perceptual system. While Appalachian English speakers may identify with, use, and more regularly hear the monophthongal form, diphthongal /aɪ/'s social connotation, which is the result of episodic encodings of its usage by certain types of speakers in particular registers via socioacoustic-encoding, results in its privileging in the perceptual system. As my

⁶ Note, the concept of Standard American English is fallible. Every English speaker speaks a unique dialect, but the perceived prestigious, unmarked standard is thought to use the unmarked diphthongal [aɪ].

results indicate, this privileging leads to the secondary perception of a glide that is absent from the speech stream.

While privilege may be present across broader linguistic contexts due to social-weighting in the speech processing system, the academic register of this study may have also encouraged Appalachian English speakers to invoke their most prescriptive language forms. Rather than (or perhaps in conjunction with) subconscious, processing bias, Appalachian English participants may have consciously accessed what they assumed to be their most “correct” underlying forms.⁷ Therefore, we see the potential for unintended effects wrought by this experiment’s procedure and presentation. For instance, choosing between options like “pie-a” and “pie-ya” likely induced reading-specific effects. Ziegler, Muneauz, and Grainger (2003) have suggested that there are orthographic effects on speech processing; an orthographic form can restructure phonological representation. The prescriptivist conventions learned by English readers and writers likely shape their processing of spoken language. Meta-linguistic commentary about standards might encourage the usage of exemplars built from experiences of careful, hyper-articulation to be used in formal disambiguation contexts (Sumner et al. 2014).

5.3. The [-ja] suffix conditions

Type 2 (control) stimuli in the [-ja] suffix condition behaved as expected. Participants overwhelmingly correctly identified [ə-ja] stimuli as the “-ya” form. While I did not have a clear hypothesis for their treatment by participants, type 3 stimuli in the [-ja] suffix condition were perceived similarly to type 1 stimuli in the [-ja] suffix condition. They were both frequently reported as being affixed with the [-a] suffix.

As shown in Figures (4a) and (4b), type 1 stimuli in the [-a] and [-ja] conditions were acoustically similar. Perhaps understanding the inherent formant change necessary in type 1 sequences, participants “factored out” an intentional [j], attributing it to mere vowel transitions.

Our type 3 stimuli, shown in Figures (6a) and (6b), were radically different across suffix conditions. Form confusability cannot be used to explain [-a] responses in type 3 [-ja] conditions. Instead, I propose that participants mapped the beginning of the suffixed [-ja] in my type 3 stimuli, which has an unavoidable F1 fall, as part of the preceding vowel. Essentially, the transition from a monophthongal [aɪ] to the high front starting place of an intentional [j] was perceived as movement to a secondary vowel target region and not the beginning of a consonant; monophthongal [aɪ-ja] was sometimes perceived as diphthongal [aɪ-a]. This result can also be explained by the social weighting of the diphthongal /aɪ/ form.

⁷ Anecdotally, multiple Appalachian English participants expressed hope that they had “gotten a good score” after completing the study or expressed misgivings and worry that their answers were “wrong.”

5.4. *Parallel Processes in Speech Perception*

The perception of phantom glides may mechanically parallel processes of phonemic restoration. Phonemic restoration (Warren et al. 1972) occurs when top-down lexical cues allow for the perceptual “filling-in” of sounds that are absent from a speech sequence. Such restoration is important for the maintenance of auditory perception in otherwise noisy and distracting environments. Hypothesized as a consequence of top-down expectancies from a lexical level, these effective restorations make test words with a consonant that has been replaced with static almost imperceptibly different from the fully articulated word (Samuel 1981). While the processes of phonemic restoration and the mapping to an erroneous underlying vowel in my test sequences lead to the similar perception of unattested sounds, the processes’ environments are quite different. No stimulus in my experiment was superimposed with static or noise; the acoustics were always continuous. Rather than “filling-in” a glide that was supposed to be in the sound sequence but was cut out or masked, participants mapped a fully articulated vowel to a different form, enabling the downstream perception of a glide.

6. **Future Directions**

To address potential confounds and shortcomings of this research, I plan to run another forced choice perception task. Because it is necessary for participants to recognize the English word being affixed, the current study used standard orthography of the words for response choices. In my follow-up study, I will use similar stimuli (types 1, 2, and 3 across both suffix conditions), but the test words will be embedded in spoken sentences in order to circumvent an academic, orthographic mapping bias. Participants will still be asked to indicate the test word’s perceived suffix, but they will select from standard “-a” or “-ya” options. Because the word will be recognizable from the context of the sentence, the words will not need to be presented orthographically. Example sentences include: “My grandmother baked a *pie*[-a/ja] for dessert” and “The heavy winds caused the *tree*[-a/ja] to fall.” Additionally, these larger sentences will serve as an opportunity for priming as the full sentence will be recorded in an Appalachian English dialect. Such attempted priming will help us better understand the robusticity of a perceptually privileged diphthongal /aɪ/. In a more vernacular, Appalachian English context, Appalachian English speakers (and perhaps speakers of other American dialects) may correctly map monophthongal [aɪ] to a low front underlying form, perceiving fewer phantom glides. Conversely, if phantom glides are perceived in Appalachian English contexts, the social-weighting of diphthongal /aɪ/ may be thought to fundamentally shape participants’ perceptual systems.

The follow-up study will also include additional survey questions to better understand individual participant's language ideologies and connectedness to Appalachian identity. As Reed's (2018) exploration of rootedness and the use of Appalachian English

features indicated, attitudes toward Appalachia and a participant's current connection to Appalachian identity affects usage of features like monophthongal [aɪ]. While they may still categorize themselves as native Appalachians, native Appalachian English speakers may be more or less likely to identify with and map to Appalachian English forms based on their regional attitudes and ongoing construction of identity. For some participants, positive ideologies about Appalachia and a concurrent privileging of Appalachian culture as an integral piece of their identity may lead to the perception of fewer phantom glides. An attachment to Appalachian English forms may subvert the socially-weighted pull of diphthongal /aɪ/ and lead to the more frequent mapping of monophthongal [aɪ] to monophthongal /aɪ/.

7. Conclusion

Participants across dialect groups perceived a consonant that was truly absent from the speech stream. Processes of perceptual mapping in my type 3 [-a] suffix condition eclipsed cues from the stimuli's continuous acoustics, incurring the perception of phantom glides. As speakers did not hear phantom glides with overwhelming frequency, I reasonably assume they otherwise mapped monophthongal [aɪ] to a low front underlying form. Suggesting an imbalance in the perceptual saliency of the diphthongal and monophthongal /aɪ/ forms, participants privileged an unfounded mapping to diphthongal /aɪ/ instead of a mapping to the low front monophthongal form they also appeared to possess. Diphthongal /aɪ/ is marked as overtly more prestigious through metalinguistic commentary, academic instantiation, and mass media. Storing such qualitative components of the form through socially-weighted encoding (Sumner et. al 2014), listeners may develop conscious and subconscious perceptual bias for diphthongal /aɪ/ and other prestigious forms. As a novel exploration of vowel mapping via the uninvestigated, secondary effect of glide perception, my findings are indicative of the subtle but powerful ways linguistic bias is perpetuated at a perceptual level.

Appendix A

List of all stimulus words by type.

Type 1: spree, glee, knee, please, tree, ski, flea, bee, key, tee

Type 2: sofa, soda, tuna, fella, llama, sauna, cola coma, puma, villa

Type 3: spy, pie, cry, dry, try, sky, buy, lie, tie, fly

Appendix B

Demographic Questions (questions without answer choices were free response)

1. Age
2. What is your gender identity?
 - a. Cisgender man
 - b. Cisgender woman
 - c. Trans man
 - d. Trans woman
 - e. Nonbinary
 - f. Agender
 - g. Genderqueer
 - h. Prefer not to answer
3. Where did you grow up?
4. Where are your parents from?
5. Where are you living now?
6. Are you a native Appalachian?
 - a. Yes
 - b. No

Appendix C

Raw Cross-Dialectal Responses

Stimuli Type	Suffix Condition	Response		Grand Total
		“-a”	“-ya”	
1 [i]-final <i>acoustic glides</i>	[-a]	264	179	443
	[-ja]	255	222	447
2 [ə]-final <i>controls</i>	[-a]	429	20	449
	[-ja]	12	459	471
3 mono. [aɪ]-final <i>phantom glides</i>	[-a]	397	62	459
	[-ja]	253	208	461
		= 1610	= 1150	= 2760

References

- Bernstein, Cynthia. 2006. Drawing out the /ai/. *Language Variation and Change in the Midland: A New Look at 'Heartland' English*. 209-232.
- Boersma, Paul & David Weenink. 2019. Praat: Doing phonetics by computer [Computer program]. Version 6.1.05. Retrieved from <http://www.praat.org/>
- Burgdorf, Dan C. & Sam Tilsen. 2021. Glides prioritize articulation, vowels prioritize acoustics. Talk given at the 2021 Annual Meeting of the Linguistic Society of America. January 7-10.
- Carver, Craig M. 1987. *American regional dialects: a word geography*. University of Michigan Press.
- Casali, Roderic F. 2011. Hiatus Resolution. In Marc van Oostendorp, Colin J. Ewen, Hume & Keren Rice (eds), *The Blackwell Companion to Phonology*. Blackwell Publishing. <https://doi.org/10.1002/9781444335262.wbctp0061>.
- Clear, Abram & Anya Hogoboom. In prep. Phonological and acoustic glides Ms. William & Mary.
- Clopper, Cynthia G., and David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32(1). 111-140. [https://doi.org/10.1016/s0095-4470\(03\)00009-3](https://doi.org/10.1016/s0095-4470(03)00009-3).
- Clopper, Cynthia G., and David B. Pisoni. 2007. Free classification of regional dialects of American English. *Journal of Phonetics* 35(3). 421-438. <https://doi.org/10.1016/j.wocn.2006.06.001>.
- Cutler, Anne, Frank Eisner, James M. McQueen, & Dennis Norris. 2010. How abstract phonemic categories are necessary for coping with speaker-related variation. *Laboratory phonology* 10. 91-11.
- Dutoit, T, V. Pagel, N. Pierret, F. Bataille & O. van der Vreken. 1996. The MBROLA project: Towards a set of high-quality speech synthesizers free of use for noncommercial purposes. In Proceedings International Congress of Spoken Language Processing, vol. 3, 1393–6. Philadelphia.
- Drummond, Alex. 2016. Ibexfarm. <https://github.com/addrummond/ibexfarm>.
- Fridland, Valerie & Tyler Kendall. 2011. Exploring the relationship between production and perception in the mid front vowels of U.S. English. *Lingua* 122. 779-793. <https://doi.org/10.1016/j.lingua.2011.12.007>
- Goldinger, Stephen. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological review* 105(2). 251.
- Goldinger, Stephen. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(5). 1166–1183. <https://doi.org/10.1037/0033-295X.105.2.251>
- Greene, Rebecca. 2010. Language, ideology, and identity in rural eastern Kentucky. Doctoral dissertation. Stanford University. <https://purl.stanford.edu/fh361zh5489>

- Hay, Jennifer, Aaron Nolan, & Katie Drager. 2006. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review* 23. 351-379.
<https://doi.org/10.1515/TLR.2006.014>
- Hazen, Kirk & Sarah Hamilton. 2008. A dialect turned inside out: Migration and the Appalachian diaspora. *Journal of English Linguistics* 36(2). 105-128.
<https://doi.org/10.1177/0075424208317127>.
- Hazen, Kirk. 2020. *Appalachian Englishes in the Twenty-first Century*. West Virginia University Press.
- Hogoboom, Anya. 2020. Phonological vs. acoustic glides. Poster given at the Berkeley Linguistics Society Workshop on Phonological Representations. University of California, Berkeley, February 7-8.
- Irons, Terry. 2007. On the Southern Shift in Appalachian English. *University of Pennsylvania Working Papers in Linguistics* 13(2). 10.
- Johnson, Keith. 2007. Decisions and mechanisms in exemplar-based phonology. *Experimental approaches to phonology*. 25-40.
- Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson, & J. Mullennix (eds.), *Talker variability in speech processing*. 145–165. San Diego: Academic Press.
- Johnson, Keith. 2005. Decisions and mechanisms in exemplar-based phonology. *UC Berkeley Phonology Lab Annual Report (2005)*. 289-311.
- Labov, William, Sharon Ash, & Charles Boberg. 2005. *The Atlas of North American English: Phonetics, Phonology and Sound Change*.
<https://doi.org/10.1515/9783v110167467>.
- Lavoie, Lisa. 2002. Subphonemic and suballophonic consonant variation: The role of the phoneme inventory. *ZAS papers in linguistics* 28. 39-54.
<https://doi.org/10.21248/zaspil.28.2002.158>
- McLennan, Conor, T. Paul Luce, & Jan Charles-Luce. 2003. Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29. 529–553.
<https://doi.org/10.1037/0278-7393.29.4.539>
- Niedzielski, Nancy & Dennis Preston. 2003. *Folk Linguistics*. Berlin and New York: Mouton de Gruyter.
- Padgett, Jaye. 2008. Glides, vowels, and features. *Lingua* 118(12). 1937-1955.
<https://doi.org/10.1016/j.lingua.2007.10.002>.
- Pierrehumbert, Janet. 2006. The next toolkit. *Journal of Phonetics* 34. 516-530.
- Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (eds), *Frequency effects and the emergence of linguistic structure*. 137-157. Amsterdam: John Benjamins.
- Pierrehumbert, Janet. 2000. The phonetic grounding of phonology. *Bulletin de la communication parlée* 5. 7-23.

- Pickett, J M & Irwin Pollack. 1963. Intelligibility of excerpts from fluent speech: effects of rate of utterance and duration of the excerpt. *Language and Speech* 6(3). 151-164.
- Preston, Dennis. 1989. *Perceptual Dialectology*. Dordrecht: Foris.
- Preston, Dennis. 1993. Folk dialectology. *American Dialect Research*. 333-378.
- Ranbom, Larissa & Cynthia Connine. 2007. Lexical representation of phonological variation. *Journal of Memory and Language* 57. 273–298.
<https://doi.org/10.1016/j.jml.2007.04.001>
- Reed, Paul E. 2014. Inter-and intra-generational /ai/ monophthongization, indexicality, and Southern Appalachian identity. *Southern Journal of Linguistics* 38(1). 159-194.
- Reed, Paul E. 2016. Sounding Appalachian: /ai/ monophthongization, rising pitch accents, and rootedness. Doctoral dissertation. University of South Carolina.
<https://scholarcommons.sc.edu/etd/3563>
- Reed, Paul E. 2018. The importance of Appalachian identity: A case study in rootedness. *American Speech: A Quarterly of Linguistic Usage* 93(3-4). 409-424.
- Samuel, Arthur. 1981. Phonemic restoration: insights from a new methodology. *Journal of Experimental Psychology: General* 110(4). 474.
- Samuel, Arthur & Saioa Larraza. 2015. Does listening to non-native speech impair speech perception? *Journal of Memory and Language* 81. 51-71.
<http://dx.doi.org/10.1016/j.jml.2015.01.003>
- Stollenwerk, Debra. 1986. Word frequency and dialect borrowing. *Ohio State University Working Papers in Linguistics (1986)*. 133-141.
- Strand, Elizabeth A. 2000. Gender stereotype effects in speech processing. Doctoral dissertation. Ohio State University.
- Sumner, Meghan. 2011. The role of variation in the perception of accented speech. *Cognition*, 119(1), 131-136.
<https://doi.org/10.1016/j.cognition.2010.10.018>
- Sumner, Meghan & Arthur G. Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60(4). 487-501.
<https://doi.org/10.1016/j.jml.2009.01.001>
- Sumner, Meghan, & Reiko Kataoka. 2013. Effects of phonetically-cued talker variation on semantic encoding. *The Journal of the Acoustical Society of America* 134(6). EL485–EL491. <https://doi.org/10.1121/1.4826151>
- Sumner, Meghan, Seung Kyung Kim, Ed King, and Kevin B. McGowan. 2014. The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology* 4. 1015. <https://doi.org/10.3389/fpsyg.2013.01015>
- Thomas, Erik. 2001. An acoustic analysis of vowel variation in New World English. *Publication of the American Dialect Society*. Durham: Duke University Press.
- Uffman, Christian. 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29(1). 451-476. <https://doi.org/10.1016/j.langsci.2006.12.017>.
- Warren, Richard M., Charles J. Obusek, & John M. Ackroff. Auditory induction: perceptual synthesis of absent sounds. *Science* 176.4039. 1149-1151.

- Williams, John Alexander. 2002. *Appalachia: A History*. Chapel Hill: University of North Carolina Press.
- Wolfram, Walt. 1977. On the linguistic study of Appalachian speech. *Appalachian Journal* 5(1). 92-102.
- Wolfram, Walt & Donna Christian. 1976. *Appalachian Speech*. Arlington: Center for Applied Linguistics.
- Ziegler, Johannes C., Mathilde Muneaux, & Jonathan Grainger. 2003. Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language* 48. 779–793.
[https://doi.org/10.1016/S0749-596X\(03\)00006-8](https://doi.org/10.1016/S0749-596X(03)00006-8)