The Construction of Social Meaning:

A Matched-Guise Investigation of the California Vowel Shift

DANIEL JAMES VILLARREAL

B.A. (College of William and Mary) 2010

M.A. (University of California, Davis) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Robert Bayley, Chair

_____

Vaidehi Ramanathan

_____

Valerie Fridland

_____

Santiago Barreda

Committee in Charge

2016

Abstract

Research on social meaning, which links language variation to the wider social world, often bases claims about the social meanings of linguistic forms on production (i.e., speakers' situational use of meaningful forms). In the case of the California Vowel Shift (CVS), an ongoing restructuring of the vowel system of California English that takes place below the level of conscious awareness, previous production research has suggested that the CVS carries social meanings of carefreeness, femininity, and privilege. Left unclear in these production-based claims is whether listeners actually pick up on and recognize the social meanings that speakers apparently utilize the CVS to transmit. In this research, a dialect recognition task with matched guises (California-shifted vs. conservative) forms the basis for exploring Californian listeners' reactions to the CVS, and how these reactions are mediated by perceptions of dialect geography. In short, this research focuses on listeners' reactions to the CVS in order to address a more fundamental question: How do listeners and speakers together participate in the construction of social meaning?

Stimuli for the main study task were drawn from excerpts of sociolinguistic interviews with 12 lifelong California English speakers from three regions of the state: the San Francisco Bay Area, Lower Central Valley, and Southern California. Guises were created from interview excerpts by modifying the F2 of each TRAP and GOOSE token via source-filter resynthesis methods. Californian guises featured backed TRAP and fronted GOOSE; conservative guises featured fronted TRAP and backed GOOSE. Ninety-seven Californians participated in a perceptual task in which they attempted to identify speakers' regional origin and rated speakers on affective scales.

The results indicated that Californians recognize the CVS as Californian, as California-shifted guises were less likely to be identified as from outside California (but more likely to be identified as from Southern California). Listeners rated California-shifted guises higher on the scales *Californian*, *sounds like a Valley girl*, and *confident*, indicating a core of social meanings indexed by the CVS. Among listeners from the San Francisco Bay Area, the CVS indexes masculinity, but among Southern California listeners, the CVS indexes femininity. Listeners from across California also rated speakers who they believed to be from the same region as them higher on *Californian*, *familiar*, and *sounds like me*.

This research demonstrates that the social meanings of linguistic forms do not reside only in speakers' situational use of these forms, as listeners did not associate the CVS with carefreeness, femininity, or privilege, the social meanings of the CVS suggested by previous studies of California English production; instead, I propose an account of the *indexical field* that links perception and production by placing the core social meanings of the CVS uncovered by this research (Californian identity, sounding like a Valley girl, and confidence) at the center of the CVS's indexical field. This research also contributes to theory in perceptual dialectology and language change. In order to explain this study's finding that the CVS is associated with Southern California, this research introduces the perceptual-dialectological process of *centrality*: the identification of speakers who are believed to most exemplify the speech of a given region. Finally, this research suggests an attitudinal stance that allows changes from below such as the CVS to flourish: speakers are aware of the change in the community (at a tacit level, if not consciously) but do not believe that they are participating in the change.

Acknowledgments

Writing a dissertation is hard.[1] Writing a dissertation completely on one's own, without the help of anyone else, is probably impossible. Adequately expressing the depth of my gratitude for those countless mentors, colleagues, friends, and loved ones who have been part of the unbroken web of support that has sustained me through six years at UC Davis? Certainly impossible. But I'll try all the same to recognize those who richly deserve it, because without all of them I would not be here.

My dissertation committee has been an incredible source of guidance throughout this project. I couldn't have asked for a better advisor than Bob Bayley. From day one of my time at UC Davis, Bob has believed in my potential and given me every opportunity to grow and thrive as a scholar. Always in my corner, Bob has been a tireless advocate every step of the way, and his encouragement and his example have been a constant source of motivation through my entire graduate career. Vai Ramanathan somehow manages an impossible balance as a mentor to graduate students: keeping her students centered and grounded while still animating them to present and publish. Vai has been invaluable to me in showing me that success as an academic needn't come at a price of one's humanity. Valerie Fridland, as a thorough, perceptive reviewer, has never hesitated to push me to do my absolute best work. The extent to which this dissertation is a cohesive work from top to bottom—and indeed, the extent to which the final chapter does more than simply summarize my results—is a testament to Valerie's tremendous influence. Santiago Barreda's patience and thoroughness helped make both the acoustic manipulation and

[1] You heard it here first!

data analysis parts of this research possible. Santiago has been generous enough to give me his time as a junior faculty member even before he had the recognition of sitting on my committee.

I also owe a debt of gratitude to many scholars who helped me in ways large and small through this project. Anne Charity Hudley, my advisor at William and Mary, has quite literally gone beyond the call of duty in continuing to guide me during my graduate career. I remain in awe of her ambition, foresight, and commitment to improving people's lives both in and out of the classroom. Michael Shepherd has been an enthusiastic advocate and a tremendous resource for recruiting participants. Cynthia Clopper generously granted me access to recordings from the Nationwide Speech Project corpus, which greatly expedited my pilot study. Bob Kennedy helped me develop my thoughts on the project and aided in recruitment. Dave Corina, Patrick Farrell, Natalie Operstein, and Andrew Wong supported my recruitment efforts. John Rickford expressed enthusiasm for my research even in its early stages. Rebecca Ambrose and Lenora Timm helped to hone my research as members of my qualifying committee. In addition, the Graduate Group in Linguistics and Office of Graduate Studies at UC Davis deserve thanks for their financial support of this research.

I can't say enough about the many colleagues, friends, and students who have left their mark on this dissertation. They patiently helped me pretest tasks and stimuli (Michaela Baer, Michelle Cohn, Julia Doyle, Renee Kemp, Emily Moline, Miranda Morris, Caitlin Tierney), generously lent their time to help with programming and/or transcription (Eric Olsen, Chris Graham, Julie Ngo), made me feel at home during my often stressful interviewing trips (Theresa and Alan Corcoran, Hakam Ghanim, Chiara McKenney, Peachy Najera, the Paulsen family), and helped me recruit speakers and listeners (Michelle Brown, Kelly Corcoran, Kristen Greer, Laurie Lawyer, Daniel Moglen, Christian Paulsen, Tracy Quan, Alissa Reyes, Syd Salsman, Sara

Thomas). Of course, this is not to mention the million intangible ways in which my friends have kept me sane, kept me happy, and kept me going.

I must finally acknowledge the contributions of loved ones, who have sustained me throughout my graduate career. Diana Villarreal and Andrew Bruno have cheered me on every step of the way and kept me grounded. Christina Calvillo has supported me through a dozen mini crises during my research, patiently held my hand whenever I was tempted toward a bleak outlook, helped me rework and clarify the most fundamental ideas behind my writing, and read every word of this dissertation (including—no joke—finding a typo in an earlier version of this sentence). My parents, James and Janet Villarreal, have always believed in me, even when the task seemed impossible; they have let me take my own path, even when they knew it was perilous; they have caught me when I stumbled, letting me be my own person but never letting me fall; and above all, they have been an endless, ever-present source of encouragement, motivation, inspiration, compassion, and love. Without my parents, I would not be here. I dedicate this dissertation to them.

Table of Contents

Table of Figures

Table of Tables

Chapter 1
Introduction

The concept of social meaning allows sociolinguists to understand how language variation links to the wider social world. The links between linguistic forms and social meanings are akin to the links between linguistic forms and referential meanings—they are almost always arbitrary. Just as no linguistic properties of the English word *cat* inherently connect to the concept of a furry feline, no phonetic properties of the *–in* ending of *workin*, *runnin*, etc., inherently connect to working-class identity. As such, social meaning must come from somewhere—it must be *constructed* somehow—but the question of how social meaning is constructed remains unanswered in sociolinguistics. Most research on the social meanings of linguistic forms focuses on how meanings are negotiated through use, implicitly suggesting that social meanings are determined to a large degree by use. This research thus focuses on listeners' reactions to variation in order to address a more fundamental question: How do listeners and speakers together participate in the construction of social meaning?

The present research centers on California English, which I argue provides an ideal setting for the pursuit of a more comprehensive picture of social meaning for several reasons. First, California English has been studied from the perspective of production and discourses but has received little attention in terms of listeners' reactions, allowing for a consideration of how these perspectives on social meaning complement one another. Second, a defining feature of California English, the California Vowel Shift (CVS), is an example of change from below—emerging outside of the conscious awareness of its users—leaving open the question of whether speech communities can assign evaluations to variation of which they are apparently unaware. Third, while California speech exists as a robust cultural object outside of the state, as demonstrated by the circulation of persistent folk-linguistic stereotypes, evidence from perceptual dialectology research indicates that Californians tend to resist these stereotypes (Bucholtz et al. 2007), calling into question how discourses about variation affect reactions to variation.

The present research pursued these questions via two experiments, a pilot study and a main study, that drew on two research strands in perceptual sociolinguistics: perceptual dialectology, which investigates speakers' mental maps of dialect variation over geographical space, and language attitudes, which investigates listeners' reactions to languages, varieties, and features. In these tasks, listeners from across California heard samples of speakers from different regions of the state, identified the regional origin of the speaker, and rated the speaker on affective scales. In the main study, each speech sample belonged to one of two matched guises differing only by the speaker's use (or non-use) of California-shifted vowels. The analysis of listeners' responses revealed crucial differences between the social meanings of the California Vowel Shift indicated by listeners' reactions—associations with California, sounding like a

"Valley girl," and confidence—versus those indicated by production or discourses, which suggest that the CVS indexes carefree privilege. I argue that this finding demonstrates that social meaning is co-constructed by speakers and listeners by creating linkages within the indexical field (Eckert 2008a).

The remainder of this chapter contextualizes the present research. I first review the sociolinguistic setting (1.2. Human geography of California) and previous dialectological research on California English, especially the California Vowel Shift (1.3. English language variation in California). I then lay out the theoretical contributions of several research traditions that bear on social meaning, especially how perceptions of language relate to perceptions of geographical space (1.4. Folk linguistics), and how the dual considerations of production and perception converge to create recognizable, circulatable social objects (1.5. Production, perception, and the creation of social meaning). With that theoretical background, I return to California to consider the status of California speech as a social object and discuss the gaps in our knowledge thereof (1.6. Visions of California). Finally, I present my research questions and outline the rest of this dissertation (1.7. The present research).

1.1. Notation

The linguistic features discussed herein primarily pertain to vowel shifts. For this reason, I have adopted John Wells' (1982) notation for lexical sets (Table 1.1), rather than the relatively static articulatory locations implied by International Phonetic Alphabet symbols. The benefits of this system are threefold. First, it does not presuppose any particular theory of the structure of American English vowel phonology (as the glide-based notation of Labov et al. 2006b does). Second, it avoids awkward statements such as "California English /ɪ/ has moved into the space of /ɛ/." Third, these lexical set names can be used in comparison to the 'ideal' vowel phonemes

represented by the International Phonetic Alphabet, either explicitly (e.g., "California English KIT has moved into the space of /ɛ/") or implicitly (e.g., "Californian English speakers front GOOSE [relative to /u/]").

Wells	IPA	Labov	Wells	IPA	Labov
FLEECE	/i/	iy	GOOSE	/u/	uw
KIT	/ɪ/	i	FOOT	/ʊ/	u
FACE	/e/	ey	GOAT	/o/	ow
DRESS	/ɛ/	e	STRUT	/ʌ/	ʌ
TRAP	/æ/	ae	THOUGHT	/ɔ/	oh
PRICE	/aj/	ay	LOT	/ɑ/	o
MOUTH	/aw/	aw	CHOICE	/ɔj/	oy

Table 1.1. Wells lexical set notation compared to International Phonetic Alphabet (IPA) and Labov (esp. Labov et al. 2006b) notation.

In addition, I have augmented Wells' notation to account for phonetic environments, since some of the features I discuss are environment-specific.[1] Environment-specific features are indicated by one or two letters affixed to the lexical set name; for instance, the "*pin–pen* merger" becomes the "KIT-N/DRESS-N merger." These notations are not intended to indicate separate phonemic classes, but are merely convenient shorthand for phonetic conditioning. These additions are detailed in Table 1.2.

[1] This idea is borrowed from Cory Holland (2014).

Affix	Environment	Example
-R	before /ɹ/	TRAP-R (*marry*)
-L	before /l/	GOAT-L (*goal*)
-N	before nasals	DRESS-N (*pen*)
-G	before velars	DRESS-G (*peg*)
-NG	before /ŋ/	KIT-NG (*thing*)
T-	after coronals	T-GOOSE (*two*)

Table 1.2. Phonetic environment-specific additions to lexical set notation.

1.2. Human geography of California

Before delving into the sociolinguistic facts of California English, it is worth briefly discussing the human geography over which this California language variation is overlaid. California is by far the most populous state in the United States, as its population of 38,802,500 represents 12.2% of the entire US population.[2] As illustrated by Figures 1.1 and 1.2, California's population is unevenly distributed, with densely populated urban areas contrasting with wide stretches of thinly populated terrain (mountainous land, forest, desert, and farmland).

[2] All population figures in this section are based on US Census Bureau July 1, 2014, population estimates, available at http://www.census.gov/popest/data/state/totals/2014/index.html (state data) and http://www.census.gov/popest/data/metro/totals/2014/index.html (metropolitan/combined statistical area data).

Figure 1.1. California relief map with place names, major highways (red), urban areas (yellow), and forest (green).
Public domain map available at https://commons.wikimedia.org/wiki/File:California_ref_2001.jpg.

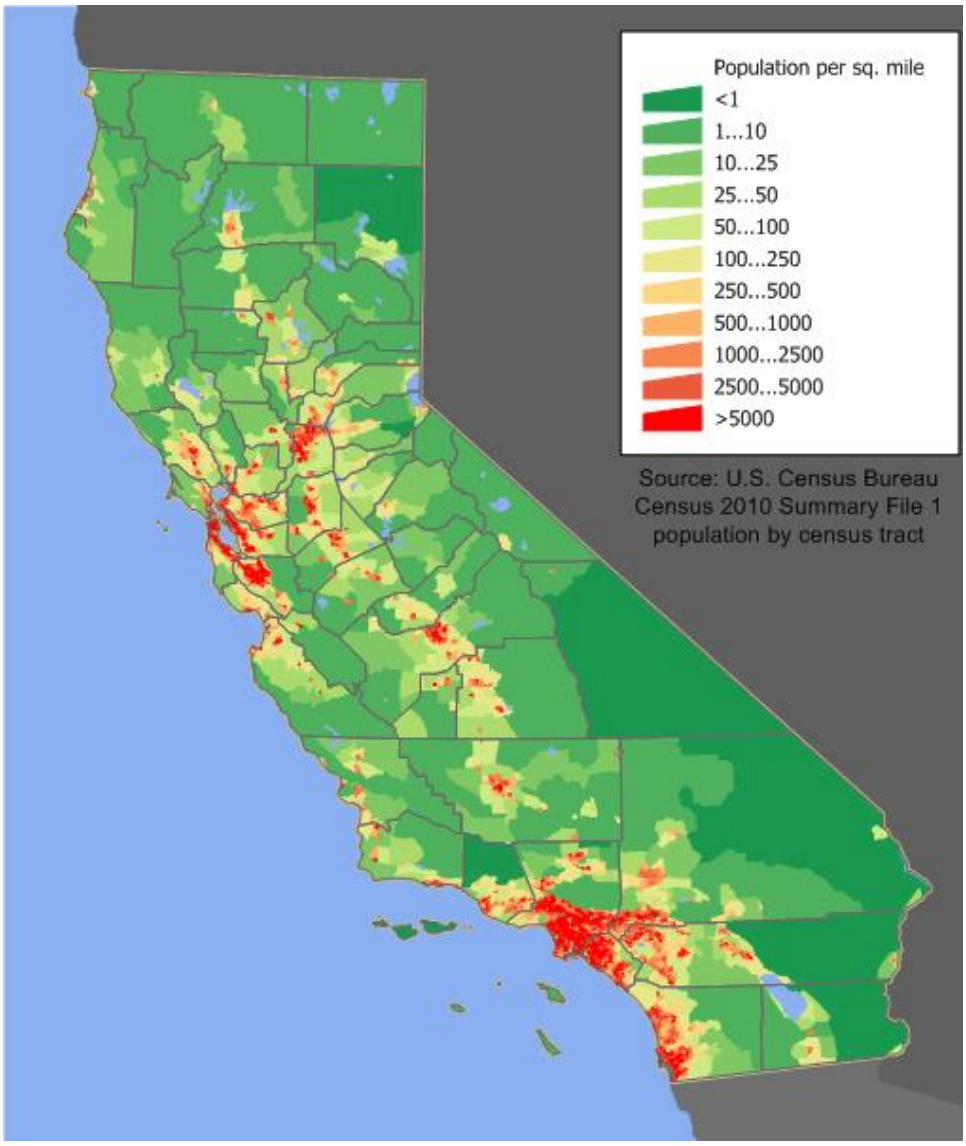


Figure 1.2. California population density by census tract.
Adapted from map at https://commons.wikimedia.org/wiki/File:California_population_map.png. Attribution: JimIrwin at the English language Wikipedia.

California's two most populous areas are Greater Los Angeles in the south (population 18,550,288) and the San Francisco Bay Area in the mid-north (population 8,607,423), which together account for 70% of the state's population. These two metropolitan areas, and the different lifestyles that each area supposedly represents, anchor a well-established shared mental representation of California human geography that divides the state along a north–south axis into "NorCal" and "SoCal," with a dividing line running east–west roughly south of Fresno (Montello

et al. 2014).[3] Less salient in dividing California's population is the coastal–inland axis (the state's curved coastline makes it hard to speak of this axis as east–west), as both Greater Los Angeles and the San Francisco Bay Area are coastal areas. The majority of inland Californian population is in the Central Valley, a vast agricultural region dotted with numerous metropolitan areas: Bakersfield (metropolitan area population 874,589) and Fresno (1,120,522) in the south, Sacramento (2,513,103) in the center, Redding (179,804) in the north, and others in between.

In later chapters, I discuss how these interregional differences formed the basis for the sociolinguistic regions used in this study's pilot study and main task.

1.3. English language variation in California

Traditional dialectological accounts place California within a larger Western region encompassing all of the contiguous US states from the Rockies westward (e.g., Clopper & Pisoni 2006). In contrast to the more fine-grained dialectal distinctions between South, North, and Midland further East—not to mention even smaller metropolitan-specific dialects such as that of Charleston and New York City—the West is often presented as relatively unremarkable. The *Atlas of North American English* (ANAE; Labov et al. 2006b) distinguishes the phonology of the West from the rest of North American English in terms of three features: fronting of GOOSE (especially after coronals) but not GOAT, the merger of LOT and THOUGHT, and the absence of

[3] The "Northern California" vs. "Southern California" distinction is, of course, subject to ideological refiguring depending on the perceiver. Residents of rural, far northern Shasta County emphasize their distinctiveness from the urbanized Bay Area (and everything south of it) by identifying it as part of *Southern* California (Podesva et al. 2013).
In addition, although San Diego is defined as its own metropolitan area (distinct from the combined statistical area centered on Los Angeles) by the US Census Bureau, the designation "Southern California" is typically understood to include San Diego; including San Diego, the estimated 2014 population of Southern California rises to 21,813,719. In keeping with this typical understanding of "Southern California," San Diego was included in "Southern California" in both the pilot study and main study (see 2.1.1. Regions and 3.1. Regions, below).

pre-voiceless Canadian raising (i.e., raising the nuclei of PRICE and MOUTH).[4] ANAE also mentions prelateral mergers between tense–lax pairs (e.g., *feel–fill*, *pool–pull*) as a possible Western feature, which accords with studies finding these mergers in the Mountain West (e.g., in Utah: Baker & Bowie 2010). Although ANAE's examples of these mergers come from Provo and Albuquerque (285–6), there is evidence for some pre-lateral mergers in California, as well. Guenter (2000) found some evidence of GOOSE-L/FOOT-L and GOAT-L/STRUT-L (i.e., *pool–pull* and *hole–hull*) merging in California and Holland (2014) found evidence for a FOOT-L/GOAT-L merger.

In contrast to ANAE, a number of studies have sought to describe the features of the English spoken in California. In an early study of California language variation, Metcalf (1972) found lexical and phonological leveling in Riverside relative to the *Linguistic Atlas of the Pacific Coast* (LAPC). Lexical forms such as *couch* or *bag* (vs. *sofa* or *sack*) that were slightly dominant when the LAPC was compiled (1952–9) had become more dominant in Metcalf's surveys. Similarly, Metcalf's informants had lost distinctions between TRAP-R/FACE-R, LOT/THOUGHT, and /w–ʍ/. Although he does not use this particular label, in essence Metcalf argues that (at that particular point in time) California English had developed into somewhat of a koiné (see e.g., Kerswill 2002).

The 1980s saw the first signs of public awareness of a distinctive California way of speaking (see 1.6.3. California enregisterment, below); it is no coincidence that Hinton et al.'s (1987) collection of studies on incipient California vowel changes was titled "It's not just the Valley Girls." These studies found GOOSE, GOAT, and FOOT fronting and unrounding; front lax

[4] Of course, none of these features is unique to the West. GOOSE fronting is found in the Southeast, Philadelphia, and Canada; the LOT/THOUGHT merger is found in Boston, Western Pennsylvania, the North Central, and Canada; and resistance to Canadian raising is found in the Midland (Labov et al. 2006b:101–14, 46).

vowels raising before nasals (as well as DRESS raising before velars) but lowering elsewhere; and THOUGHT lowering. The most general of these changes were later formalized by Eckert (2008b) as the *California Vowel Shift* (CVS). Eckert's diagram of the shift is reproduced with lexical set notation as Figure 1.3. Here, the merged LOT/THOUGHT phoneme occupies the higher and backer space of /ɔ/ rather than /ɑ/, although as noted below, the location of this merged phoneme is a possible locus of intrastate variation. In addition, consistent with Hinton et al.'s findings about front lax vowels, TRAP evinces a *nasal split*, raising before nasals but backing elsewhere. Eckert does not follow Hinton et al., however, in extending this nasal split to the other front lax vowels, nor does she identify other phonologically conditioned changes, such as the prelateral mergers suggested by ANAE or the prerhotic mergers noted by Metcalf.

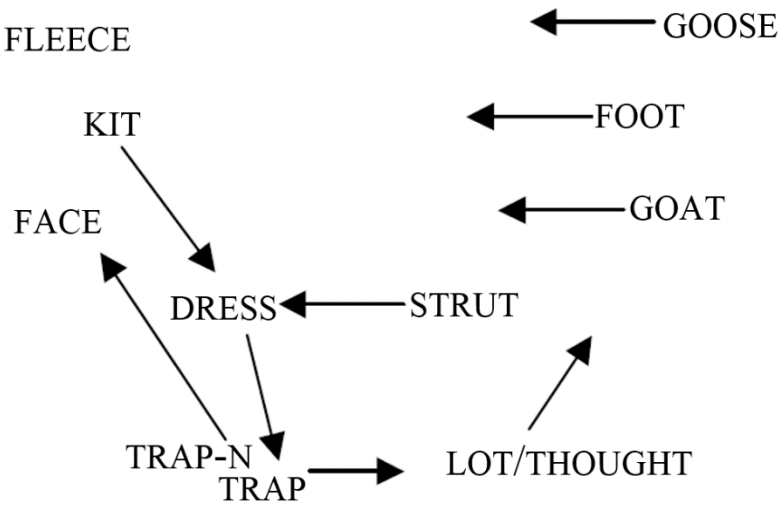


Figure 1.3. The California Vowel Shift, adapted from Eckert (2008b:34, Figure 1), with lexical set notation instead of Eckert's.

Eckert's general findings are echoed by Hagiwara (1997) and Kennedy and Grama (2012). Holland's (2014) study of speakers from the Bay Area, Southern California, and Central Valley indicates a similar, albeit somewhat more conservative, vowel space, with the mean of all speakers' normalized KIT vowels falling in the ANAE "midhigh" range, TRAP-N in the ANAE

"midlow" and "midback" ranges, and GOAT in the ANAE "midback" range. Although it is seldom described this way in the literature, the CVS is a clear example of a *change from below* (Labov 1990), as it is an innovation, rather than an importation of an external standard, led by female speakers (Cardoso et al. in press; D'Onofrio et al. in press; Holland 2014; Kennedy & Grama 2012).

It should be noted that, for the most part, the studies reported in this section based their data mostly or solely on Caucasian speakers, on the surface implying that to sound Californian is only to sound Californian and Caucasian. In reality, it is perhaps most useful to think of sound changes labeled "Californian" as resources present within a cross-ethnic linguistic marketplace (Bourdieu 1977) that happen to have mostly been utilized by Caucasians (see e.g., Fought 1999). However, as these notions rely on concepts that are out of place in a survey of dialectological research, I will turn to intrastate (interregional) variation in California English before returning to ethnicity later in this review (see 1.6.1. California indexicality, below).

1.3.1. Intrastate variation in California English

In general, earlier survey studies of California English did not consider a possible role of intrastate regional origin as an independent variable. (An exception in this regard is Metcalf, who compares LAPC data from the entire state to the "Southern Californian English" examined in his Riverside surveys.) For example, Hagiwara's speakers were all from Southern California, and Kennedy and Grama's speakers were all from coastal areas between San Francisco and San Diego. Some studies have tacitly assumed that a regionally restricted sample would stand in for the entire state. Most of Hinton et al.'s speakers came from the (San Francisco) Bay Area, although the authors note that one of their most conservative speakers was from a small town outside of the urbanized Bay Area (121). Eckert's speakers were Bay Area elementary school

students; confusingly, Eckert attributes the vowel shift outlined in Figure 1.3 to all of California in the body text of her article but to "Northern California" (i.e., the Bay Area) in both her figure caption (34) and her Stanford webpage on vowel shifts (http://www.stanford.edu/~eckert/vowels.html, accessed 1/31/2016).[5]

Recent studies, however, have begun to address potential intrastate variation, specifically with respect to CVS features. In short, the picture that emerges from these studies is one in which the CVS has advanced throughout the state, but with subtle local/regional differences. Hall-Lew's (2009) dissertation focused specifically on phonetic variation and social meaning in a single neighborhood (the Sunset District) of San Francisco. Whereas the older San Franciscans in Hall-Lew's study demonstrated resistance to the LOT/THOUGHT merger, a positive correlation of age with the extent of merger suggests that San Francisco is moving toward merger in apparent time (157). More recent work using the same corpus has found that TRAP-G patterns with TRAP-N in raising among San Francisco speakers, which is an exception to the general CVS nasal split in which TRAP-G backs along with TRAP (Cardoso et al. in press).

D'Onofrio et al.'s (in press) study of low vowels in three Central Valley cities finds that the merged LOT/THOUGHT vowel is produced in the higher and backer space of /ɔ/ rather than /ɑ/; in other words, among Central Valley speakers LOT raised to merge with THOUGHT. By contrast, coastal speakers' merged vowel tends to be produced in the lower space of /ɑ/, suggesting that

[5] Eckert's webpage also mentions a velar-nasal split for KIT such that the vowel raises and tenses before [ŋ] and lowers elsewhere. While Hinton et al. do not report KIT-NG raising, this change is related to variables that they do report: raising of KIT-N, DRESS-N, TRAP-N, and DRESS-G. Cardoso et al. (in press) also report KIT-NG raising among San Francisco females, with no apparent-time effect, although it is difficult to know whether KIT-NG is raised or simply resists the lowering and backing of KIT in other environments because the authors do not provide FLEECE as a reference point.

Informally, I have noticed this feature among numerous West Coast natives. In addition, five exceptionally clear tokens of KIT-NG raising can be heard in the first verse of the 2000 hit single "All the Small Things" by the San Diego-based band Blink-182: http://youtu.be/ahWmkV0mtvk?t=14s.

THOUGHT lowered to merge with LOT (Kennedy & Grama 2012). (As a result, the schematized CVS in Figure 1.3 actually reflects Central Valley speech to a greater extent than coastal speech.) In addition, whereas several studies identify TRAP as lowered (Cardoso et al. in press; Holland 2014; Kennedy & Grama 2012), D'Onofrio et al.'s Central Valley speakers actually *raise* TRAP in apparent time; the raising of LOT means that TRAP is still these speakers' lowest vowel.

Aiello (2009) compared vowel production by men and women from the Bay Area and Southern California, finding no significant intrastate differences in spectral properties. Aiello did find an interaction between region and gender with respect to speech rate, however. Whereas Southern California men and women had comparable speech rates, Bay Area speakers differed by gender, as Bay Area females spoke slower, and males faster, than their Southern California counterparts (18). Holland (2014) likewise found few intrastate vocalic differences; Southern California speakers had a backer DRESS, Bay Area speakers had a fronter TRAP-N, and Central Valley speakers (and some Southern California speakers) merged KIT-L and FLEECE-L.[6]

ANAE shows striking intrastate variation in the KIT-N/DRESS-N merger; whereas almost all Caucasian speakers in San Francisco, Los Angeles, and San Diego exhibit no merger, all five (Caucasian) Sacramentans are close in perception or production and four of five lower Central Valley (Fresno/Bakersfield) residents are fully merged (68, Map 9.5). More recent research, however, suggests a reversal of this merger in Bakersfield, as speakers 26 and younger produced distinct KIT-N and DRESS-N vowels (Warren & Fulop 2014); this distinctive Lower Central Valley

[6] Holland's speakers all had ties to a university near Sacramento and had lived in the Sacramento region for at least a year. As such, Holland cautions that "this study does not attempt to make claims about vowels in regions outside of the Central Valley, but instead investigates whether any differences that do exist are carried to the Central Valley by immigrants from other regions of the state" (131). It is also worth noting that the large majority of Holland's Central Valley speakers were from either Sacramento or Modesto.

feature may be leveling to match the rest of California. In addition, ANAE indicates that T-GOOSE is more fronted in Southern California than further north (101, Map 10.24). The vowel in *roof* is /ʊ/ for the Central Valley and /u/ for San Diego/San Francisco, with Los Angelenos divided between /ʊ/ and /u/ (291, Map 21.2).

As discussed in Chapter 3 below, the matched-guise stimuli in the main study were created by resynthesizing TRAP and GOOSE tokens in interview excerpts. In keeping with research on CVS production, the California-shifted guises featured backed TRAP and fronted GOOSE, whereas the conservative (non-shifted) guises featured fronted TRAP and backed GOOSE. These features were selected for manipulation because they each stand in for a major subsystem of the CVS, both have featured in popular portrayals of California English speakers, and both are used by Latina/o speakers as well as Caucasian speakers to index their place in the social landscape (see 1.6.1. California indexicality, below).

1.4. Folk linguistics

Languages, language varieties, and linguistic features are only partially defined by their production, distribution, and variation (the objective data uncovered by linguists using scientific methodologies). Most individuals attach subjective meanings, beliefs, and associations to the objects of linguistic study, and these folk-linguistic 'facts' may fail to match objective linguistic data, contradict one another, and/or contribute to deeply pernicious processes such as social subordination and denial of educational access (Labov 1972a; Lippi-Green 2011). That linguists may find these folk beliefs comical, repulsive, or anywhere in between does not mean that folk beliefs should be ignored; rather, what non-linguists believe about language has been increasingly recognized as a legitimate and even important object of serious linguistic study (Niedzielski & Preston 2003; Paveau 2011).

Accuracy is, moreover, only one dimension along which folk beliefs can be categorized. Preston (1996a:40–41) defines four "modes of folk linguistic awareness": availability, accuracy, detail, and control. *Accuracy* is the degree to which non-linguists' beliefs about language match objective linguistic data; *availability* is the degree to which non-linguists can or do comment upon topics of language; *detail* is the specificity of folk beliefs; and *control* is non-linguists' ability to perform a variety/feature. Folk beliefs may thus be conceptualized as 'settings' along (relatively) continuous dimensions that are (relatively) mutually independent.

The complexity of folk-linguistic beliefs means that linguists have brought a diverse set of research traditions to bear on the topic. The research strains described in the following sections, language attitudes (1.4.1) and perceptual dialectology (1.4.2), represent only a subset of possible tools for investigating folk linguistics (Preston 2011). Of particular import to the present research is the underexplored field of dialect recognition, which I explore in detail below (1.4.3).

1.4.1. Language attitudes

The field of language attitudes investigates the attitudes that listeners hold toward given languages, varieties, or features (Giles & Billings 2004). Language attitudes research often utilizes the matched-guise technique (MGT) pioneered by Lambert et al. (1960). The crux of an MGT is comparing listeners' reactions to pairs of stimuli (guises) that differ only in their use of given languages, varieties, or features (which are in reality produced by the same speaker). (For brevity's sake, I refer to "both" guises or "pairs" of guises, though in reality there may be more than two guises.) This method allows allow researchers to investigate the social meaning of languages, varieties, or features—the impact that using a certain feature has on how speakers are perceived—such as French vs. English use (the original Lambert et al. 1960 study), Welsh vs. English accents (Coupland et al. 1999), or [ɪn] vs. [ɪŋ] realizations of the English *–ing* morpheme

(Campbell-Kibler 2008; Loudermilk 2013). As Giles and Billings (2004:190) note, this method presents several distinct advantages, among them the ability to control for extraneous variables and the elicitation of "apparently private attitudes" that perceivers might not reveal if asked directly, given taboos against the overt expression of prejudicial opinion (e.g., the "political correctness effect" explored by Bucholtz et al. 2008:76ff). Listeners are typically asked to perform evaluations via ratings on semantic differential scales (e.g., *friendly–unfriendly*); across studies, these ratings generally cluster around *status* (e.g., *educated*) or *solidarity* (e.g., *helpful*) constructs (Edwards 1999:102).

This study involves the most common type of MGT, in which guises substantively differ only in a crucial feature, which present more complex questions for task design than language or variety-level guises. Implicit within the definition of the MGT are several criteria for valid tasks, which I label opacity, naturalness, and generalizability. Combined with considerations such as the feature under investigation and the researcher's tolerance for noise, these criteria influence choices for design features of the task such as the number of trials, the method by which guises are created, and how the task is framed for listeners.

Opacity means that listeners are unaware of the guise manipulation, such that listeners believe they are judging "(supposedly) *different speakers*" (Giles & Billings 2004:189, emphasis original). For example, consider a task with two trials whose stimuli were recorded by the same speaker differing only in the use of creaky voice. A listener who heard both trials would recognize the speaker as being the same and would have difficulty rating the guises differently on attributes such as intelligence or sexuality that are typically thought to be invariant personal qualities. In other words, opacity preserves the task's ability to uncover the social meaning of the crucial feature. As the prior example suggests, opacity may be fulfilled partially through the

number and order of trials that listeners hear. If listeners are to hear both guises, these trials must be separated by enough filler trials so the speaker in the crucial trials is not recognized. Alternatively, listeners may be assigned to hear one guise or the other; this method has the benefit of facilitating matched guises from multiple speakers but the drawback of requiring a larger sample size.

Naturalness and generalizability arise from the property that guises substantively differ only in the crucial feature. *Naturalness* means that both guises are (equally) plausible as something that the speaker could say. For example, a stimulus featuring a higher-pitched female speaker whose intonation was manipulated to produce a terminal pitch contour ending at 100 Hz would not sound plausible. Naturalness is not equivalent to representativeness, however; stimuli need not reflect the way a speaker actually speaks so long as the stimulus is plausible (i.e., something that the speaker *could* say even if it is not something they *would* say). For example, a stimulus that features several tokens of prevocalic word-final consonant cluster reduction can be plausible even if the actual speaker reduces prevocalic consonant clusters at a very low rate.

Generalizability means that each guise adequately represents one variant of the feature being studied and that the difference between guises must be perceptible at some level. For example, in a study on perceptions of discourse marker *like*, a stimulus where *like* is literally every other word would fail generalizability, as this stimulus would fail to represent speakers' actual use of *like*. In studies where matched guises are created for multiple speakers, generalizability takes on the added dimension of requiring that a certain guise be represented to an equivalent degree by all speakers. In a study of word-final (z)-devoicing with multiple speakers, for example, it would not work to have one speaker's "devoiced" guise feature 50% devoiced tokens and another speaker's "devoiced" guise feature 100% devoiced tokens.

Matched-guise studies commonly use *carrier phrases* (i.e., a stretch of speech containing one or more tokens of the crucial feature), which must either be read from a script or produced spontaneously. Carrier phrases can bolster opacity, as the feature is less obvious to listeners if embedded within a phrase, and naturalness, as an acoustically manipulated token that sounds suspect in isolation may be less noticeable in the course of a larger phrase. While spontaneous speech is more naturalistic, it lessens the control that the researcher has over stimulus content, possibly introducing extraneous features; in multi-speaker designs, this variation in stimulus content favors multiple guise conditions over a filler-trial design, as listeners are likely to notice repeated content. At first glance, the criterion of generalizability would seem to strongly favor read speech, as the researcher can isolate listeners' attitudes toward the crucial feature by using evaluatively neutral content across speakers. Campbell-Kibler (2006:82–85) argues convincingly, however, that such neutral content may not be desirable. Not only can listeners detect that speech is read rather than produced spontaneously, but this detection of a reading register can significantly alter their judgments of the speaker. It is also doubtful whether any content is ever in reality 'neutral,' since listeners' interpretation of content is filtered through that of form (Campbell-Kibler 2009)—at the same time, researchers should be wary of any content that overly influences perceptions. In addition, listening to identical content across multiple trials raises the possibility of introducing listener fatigue.

Broadly speaking, past matched-guise studies have utilized three types of methods to create different guises: complete re-recording, re-recording with splicing, and recording with acoustic manipulation. In the first method, the speaker simply records (or performs) both guises. In the second method, the speaker records the same content twice, but smaller parts of one guise

are spliced onto the other so that the two guises are identical save for the critical portions. In the third method, the speaker records just one guise, and the other guises are created from it.

No method is inherently superior to the others, but may be more or less suitable depending on the features being manipulated (especially features' level of detail and speakers' ability to control features), the context of responses, and the technology available. The complete re-recording method is most suitable for studies where the guises differ by the most global feature possible—language/dialect—and the speaker is a competent performer of both varieties (e.g., Cargile 1997; Lambert et al. 1960; Purnell et al. 1999). The splicing method is more suitable when the feature is one that speakers can control, such as (ING) (Campbell-Kibler 2006) or PRICE monophthongization (Allbritten 2011); speakers' ability to consciously control a feature, of course, depends on their being aware of it.

The final method, acoustic manipulation, is most appropriate when speakers cannot control a feature and/or when it would be difficult for speakers' re-recordings to adequately resemble the original. The main advantages of this method, then, are not only that it makes possible matched-guise studies on a wide range of variables, but also that it does not require any re-recording, so stimuli can be created from naturally occurring speech. The main disadvantage is that naturalness is difficult to achieve; if speech signals are not edited carefully, they can sound computer-like (although, as mentioned above, carrier phrases increase the 'margin for error' in this respect). For this reason, this method has often been used for guises that require only simple manipulations, such as Dailey-O'Cain's (2000) method of deleting *like* tokens to create guises with and without *like*. When vowels are manipulated, on the other hand, it is typically in the context of relatively abstract single-word (i.e., absent a larger carrier phrase) guises on a continuum between two acoustic tokens (see e.g., Fridland et al. 2004; Plichta &

Preston 2005). Graff, Labov, and Harris (1986), however, did utilize vowel resynthesis to create matched guises from the naturally occurring speech of a single speaker.

The nature of the current study, in which matched guises were created from multiple speakers' spontaneous speech by using resynthesis to embed resynthesized vowels within carrier phrases, presented unique challenges to creating stimuli that satisfied opacity, naturalness, and generalizability. I ultimately determined that acoustic manipulation was the only suitable method for creating these guises (see 3.4.1. Matched-guise study design, below); the procedures by which manipulation targets were calculated (see 3.4.2. Manipulation target calculation, below) and the manipulation was implemented (see 3.4.3. Implementation of manipulation and Appendix E, below) were also designed with the criteria of naturalness, generalizability, and opacity in mind.

1.4.2. Perceptual dialectology

The field of perceptual dialectology (PD) seeks to understand speakers' mental maps (Gould & White 1986) of dialect variation over geographical space; that is, PD is the folk linguistics analogue of regional dialectology. PD utilizes a variety of methods, some of which are borrowed from general research on linguistic perception. Map-drawing tasks, on the other hand, are unique to PD; in these tasks, participants are given a map of a region (e.g., the United States) and asked to indicate where people speak differently (i.e., by drawing borders and labeling regions). A sample map from Dennis Preston's work (1996b) is given in Figure 1.4.

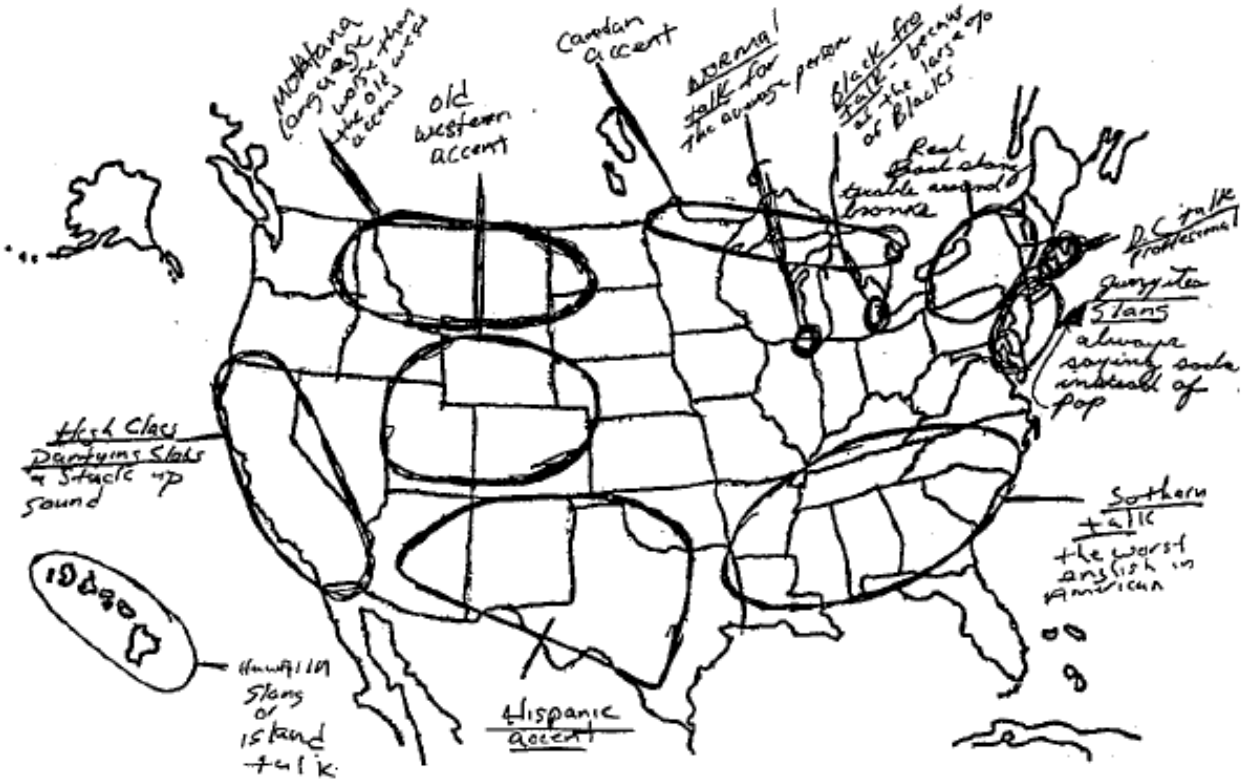


Figure 1.4. Hand-drawn perceptual dialectology map by a Chicago native (Preston 1996b:307, Figure 8). Reprinted with permission from *Focus on the USA* (ed. by Edgar W. Schneider), https://benjamins.com/#catalog/books/veaw.g16/main.

This map is clearly rich in folk-linguistic data. Despite the large number of speech regions that are salient—i.e., highly available—to this speaker (Hawai'i, Detroit, Washington, DC, and eight others), much of the map (e.g., the Pacific Northwest, the Great Plains) receives no commentary. Second, while each regional label in this map does include some metalinguistic content,[7] labels vary on the dimension of detail, with some naming particular distinctive features ("always saying soda instead of pop") but most simply indicating a global level of detail ("old Western accent"). While a "stuck up sound" has no inherent linguistic reality, of course, for this folk commentator it is sufficiently self-evident to function as a linguistic descriptor that their

[7] Labels on PD maps often lack any kind of discernable metalinguistic detail, as with the label "Potatoe [sic] Land" one respondent assigned to Idaho (Preston 1996b:308, Figure 9).

audience is assumed to readily recognize. In other words, this descriptor may not be detailed in the sense of linguistic features, but it is certainly available. Relatedly, several labels reveal attitudinal associations (which were not asked for but supplied anyway): "the worst English," "normal talk," "professional," etc. In other words, perceptions of region do not operate in a vacuum, but instead may serve as an organizing principle for other sorts of folk-linguistic beliefs (a point revisited in 1.5.4. Linguistic sense of place, below).

One robust finding across PD studies of the United States is the general salience of the South, which is by far the most frequently identified region on hand-drawn maps (Preston 1996b:305), regardless of the regional origin of the respondents (Fought 2002; Hartley 1999; Lance 1999). The South's salience is no doubt due to its unique stigmatization, as revealed by both map-drawing tasks and correctness/pleasantness tasks, in which respondents rate each state on the correctness and pleasantness of their speech. The correctness ratings elicited by Preston (1996b) demonstrate a consensus among Michigan, Indiana, and Southern raters that the least correct English is spoken in the South as well as in New York City, with the most correct English found (for the most part) in the Upper Midwest (312–13). With respect to pleasantness, however, this consensus is replaced by a preference for local speech (316–17); while Michiganders and Southerners agree that Southern speech is not correct, and Michiganders find Southern speech the least pleasant, Southerners find their own speech the *most* pleasant (and don't particularly care for Michiganders'). These correctness/pleasantness results suggest that processes central to language attitudes may also operate on speakers' perceptions of language variation over space; indeed, Preston (1999) showed that Michiganders rated Northern speech substantially higher on correctness scales (e.g., *educated, good English, no twang*) but Southern speech slightly higher on pleasantness scales (e.g., *casual, friendly, down-to-earth*).

Numerous PD studies have demonstrated the effect of regional identity on speech perception by utilizing auditory guises, often with phonetic manipulations. Plichta and Preston (2005), for example, found that listeners from Michigan and Indiana were able to distinguish PRICE tokens on a fine-grained continuum between monophthongal and diphthongal PRICE, attributing more monophthongal tokens to more Southern speakers. Such sensitivity to fine phonetic detail may be surprising in light of the relatively crude, under-detailed labels found on hand-drawn maps; in other words, distinctive features may in fact be perceptually available to speakers even in absence of overt folk-linguistic commentary.

PD studies have also demonstrated the usefulness of comparing how dialects are perceived by both *ingroup* members (speakers of that dialect) and *outgroup* members (speakers not of that dialect). Allbritten (2011) found that perceptions of Southernness and rurality strengthened with an increase in Southern features, indicating that rurality is connected to the notion "sounding Southern." Not all features created this effect uniformly, however; while the (ING) variable did not affect Southerners' perceptions of Southernness, non-Southerners perceived [ɪn] samples significantly more Southern than [ɪŋ] samples. That is, whereas [ɪn] plays into non-Southerners' percepts of Southern speech, Southerners don't necessarily view themselves as [ɪn] users (186). In other words, with respect to (ING), Southerners (ingroup perceivers) have a different notion of what it means to "sound Southern" than do non-Southerners (outgroup perceivers).[8]

[8] To that end, Labov et al. (2011) showed that Southerners did not downgrade [ɪn] on professionalism to the extent that Northerners did. Californians also apparently have an expectation that they are [ɪŋ] users. Loudermilk (2013:113–34) found that Californians encountered greater processing difficulty when hearing a Californian using [ɪn] or a Southerner using [ɪŋ] than a Californian using [ɪŋ] or a Southerner using [ɪn].

Niedzielski (1999) presented Michiganders with samples of a single Detroiter's speech, asking listeners to match the vowel heard in a given word with one of six possible vowels along a continuum (e.g., MOUTH tokens differing by the extent of nucleus raising, a Canadian raising feature also present in Detroit). These samples were presented with the label "Michigan" or "Canadian," indicating the (purported) regional identity of the speaker. When told that the speaker was Canadian, these Michiganders matched the vowel in *house* with a Canadian-raised MOUTH token, but when told that the speaker was from Michigan, listeners matched the vowel with a nonraised MOUTH. When presented with the fronted LOT vowel in the word *pop*, listeners were more likely to hear a hyper-standard vowel (further back than canonical /ɑ/) than the actual Northern-Cities-Shifted token when they believed the speaker was a Michigander. Niedzielski (2002) attributes this effect to Michiganders' strong folk belief in the standardness of their own speech; in the ears of Michiganders, a fellow Michigander couldn't possibly sound like someone with raised MOUTH and fronted LOT (never mind phonetic reality). Between Niedzielski, Plichta and Preston, and Allbritten (plus Labov et al. and Loudermilk, fn. 8), we get a complicated picture of the relationship between phonetic detail and perceptions of regional identity. In some cases, listeners are highly sensitive to phonetic detail, but in others, extralinguistic information can 'desensitize' listeners.

1.4.3. Dialect recognition

A small body of research on *dialect recognition* has investigated listeners' ability to 'place' a speaker's voice based on acoustic information. These tasks ask listeners to match a series of sound clips to speech regions and to rate speakers on various semantic differential scales. Research on dialect recognition demonstrates that it is a sociolinguistic process rather than a purely phonetic process; that is, when listeners make a judgment on where a speaker is

from, they are not simply matching the speaker to preexisting acoustic templates of regional speech, but they are also drawing on their attitudes toward certain regions, including attitudes toward their own region.

In Williams, Garrett, and Coupland's (1999) original study of dialect recognition, adolescent listeners in the six perceptual-dialectological regions of Wales listened to 30-second excerpts of personal narratives, rated speakers on affective scales (e.g., likeability, "Welshness"), and made a forced-choice guess at speakers' regional origin. The 14 speakers were adolescent boys from each of the six Welsh regions, plus two Received Pronunciation (RP)-speaking English students.

The overall rate at which adolescent listeners accurately recognized speakers' regions was roughly 25%, consistent with Clopper and Pisoni's (2004) 30% recognition rate among US regions. (Here and below, I use *accurate recognition* to refer to correctly identifying a speaker's region of origin, in contrast to *identification*, which is indifferent to accuracy.) Teachers fared much better on the recognition task, accurately recognizing the majority of speakers at a rate of 45% or greater, likely due to greater exposure to variation. The divide between teachers' and students' accurate recognition was greatest for the RP speakers, possibly reflecting the salience of RP as a prescriptive variety among gatekeepers of the standard language.

Williams et al.'s results also mirror the principle that perceptions of a dialect often differ between ingroup and outgroup perceivers (e.g., Allbritten 2011), in that speakers were accurately recognized at a greater rate by ingroup listeners (i.e., listeners from the same region as the speaker) than outgroup listeners (i.e., listeners from regions other than the speaker). This effect was not uniform, however, as some speakers were accurately recognized by ingroup listeners at a

rather low rate. The authors' analysis of likeability ratings helps to explain this variation in accurate-recognition rates, as the speakers who were rated most likeable were also more likely to be identified as belonging to a listener's ingroup, whether or not this was actually true. Cardiff listeners, for example, found Northwest speaker 2 more likeable than Northwest speaker 1 and misidentified Northwest speaker 2 (but not Northwest speaker 1) as a Cardiffian. From this pattern, the authors find evidence for a social-psychological *claiming* effect, with individuals grouping themselves with those they find desirable (though it is not clear which features lead to listeners' judgment that the speaker is desirable). The concept of claiming features prominently in my analyses of both pilot study (see 2.2.3. Interaction of identification and language attitudes, below) and main study responses (see 4.2.7. Interaction of listener region with perceived region, below).

Finally, the authors found an interesting effect relating to "Welshness" ratings; whereas regional pairs of speakers were often recognized at very different rates, they tended to cluster together on "Welshness." This result indicates that listeners "do in fact recognize (at some level of awareness) that there are pairs of speakers from the same dialect community" (354). While the identification task taps into perceptual categorization that is less available to these students, they are still able to reliably group speakers along a more available notion of centrality (or peripherality) to Welsh linguistic identity. In short, by pairing regional identification with attitudes and eliciting responses from regionally diverse samples of listeners, dialect recognition tasks are able to elicit surprisingly rich results that illustrate folk-linguistic awareness on several levels.

1.5. Production, perception, and the creation of social meaning

In the previous sections I discussed speakers of California (1.3. English language variation in California) and sociolinguistic perception (1.4. Folk linguistics) as though production and perception were wholly separate processes. In reality, the two are inextricably linked; nowhere is this more clear than in the "third wave" of variationist research, which applies an analytic focus to linguistic variables as "resource[s] for the construction of social meaning" (Eckert 2005:1). At the root of this pursuit is the question of how facts about language, which lack inherent value (Giles et al. 1974), take on the sorts of social meanings that can eventually become the subject of overt commentary by non-experts. (In other words, how did we go from the California Vowel Shift to "The Californians"?)

In the following sections, I first consider how beliefs about language come to exist in the first place (1.5.1. Language ideologies), how linguistic variables can acquire social meaning (1.5.2. Indexicality), and how these social meanings can create social objects out of language varieties (1.5.3. Enregisterment), with particular focus drawn to connections between language and geographic space (1.5.4. Linguistic sense of place).

1.5.1. Language ideologies

Silverstein's (1979:193) definition of language ideologies as "sets of beliefs about language articulated by the users as a rationalization or justification of perceived language structure and use" suggests an obvious affinity with the study of folk linguistics. Indeed, most approaches to language ideologies reject the view sometimes taken by formal linguists of the irrelevance of "secondary responses" (Woolard 1992:239). While there are multiple approaches that fall under the banner of language ideologies (Woolard 1992), I will focus on the applications of language ideologies to the types of perceptual data discussed above. These applications are

relevant since, in a sense, language ideologies are what make folk-linguistic beliefs possible. For example, if American society did not have a standard language ideology, a notion of standard language that permits the differential valuation of linguistic varieties (Lippi-Green 2011; Milroy & Milroy 1999), then we would be surprised to find American listeners rating different speakers as more or less fit for broadcasting work (Labov et al. 2011) or Michigan listeners unable to hear 'nonstandard' features in Michigan speech (Niedzielski 1999).

Irvine and Gal's (2000) trio of semiotic processes of linguistic differentiation offers coherent and powerful tools for explaining language ideologies. *Iconization* suggests some sort of naturalized linkage between linguistic features and social attributes, imbuing, in other words, linguistic objects with (differential) inherent value. *Fractal recursivity* creates an opposition on some level of relationship (social, linguistic) by relating it to an opposition on another level. Finally, *erasure* works in the opposite direction, removing degrees of complexity from the "sociolinguistic field" by denying the presence of certain groups, practices, or phenomena.

These semiotic processes allow us to ground several observations about the perceived relationships between language and place shown in PD research (see 1.4.2. Perceptual dialectology, above). For instance, the stereotypical dichotomy between US Northern vs. Southern speech (see e.g., Preston 1989) shows all three processes at work. To begin, this dichotomy is made possible in the first place by fractal recursivity, as it projects historical sociopolitical/cultural differences onto language (or folk geography onto folk dialectology).[9] Preston (2011:19) notes that this dichotomy allows for a dichotomized notion of speech rate and

[9] That (non-folk) dialectologists also find substantial differences between Northern and Southern speech is mostly beside the point. The Midland speech region identified by dialectologists is absent from folk dialectology—no doubt since, unlike the North/South distinction that is so pervasively salient to American history and culture, Americans lack a convenient folk-geographic notion of the areas the Midland defines (thus blocking the operation of fractal recursivity).

friendliness: "speakers of Southern American English are said to speak slowly because they are slow (but hospitable); speakers of Northern American English (especially East Coast urban Northern) speak fast because they are fast (and rude and inhospitable)." In other words, there is an iconized connection between speech rate and cognitive rate (i.e., intelligence); when related to the status/solidarity opposition (see 1.4.1. Language attitudes, above), this means that a fast speaker is not a friendly speaker. This dichotomy also erases Western speech, as both the sociopolitical/cultural and dialectological profiles of the West do not map cleanly onto the Eastern-based construal of North vs. South—which is not the same dichotomy as "NorCal" vs. "SoCal."

1.5.2. Indexicality

As mentioned above, third-wave variationist research focuses on the construction of social meaning via linguistic variables; this notion has been formalized as *indexicality*, "the creation of semiotic links between linguistic forms and social meanings" (Bucholtz & Hall 2005:594). Ochs's (1992) influential article addresses the question of how associations between linguistic features/practices and gender (for example, the "femaleness" of *wa* and the "maleness" of *ze* among Japanese sentence-final emphatic particles) enter the social space of a speech community—attempting, in other words, to square the "inherent" gendered value of certain linguistic practices to their availability for use by both male and female speakers. To tackle this problem, Ochs formulates the notion of *indirect indexicality*, wherein "gendered" features do not index gender directly, but instead index stances, acts, and activities that in turn index gender themselves (341–43). To return to the example of Japanese sentence-final emphatic particles, *wa* directly indexes a "delicate intensity" and *ze* a "coarse intensity," two stances that then (through the ideological relations connecting femaleness to delicacy and maleness to coarseness) come to

index femaleness and maleness, respectively. This account also helps to explain why males are free to use *wa* and women free to use *ze*; since these particles index stances, they can be used by the "marked" gender to indicate a stance that is atypically delicate (for a male user) or atypically coarse (for a female user).

Silverstein (1992) inductively extends indirect indexicality by replacing Ochs's static structure of two indexing relationships to theoretically infinite "orders of indexicality," as n^{th}-order indexicality can always give way to indexicality at order n+1 (Silverstein 2003:212). Silverstein (2003:217) further relates indexicals at orders n, n+1, and (n+1)+1 with Labov's (1972b:534–35) trichotomy of linguistic variables as indicators (unconscious variables that are straightforwardly correlated with group membership), markers (unconscious variables that show stylistic stratification), or stereotypes (variables that are the subject of overt commentary and may, as a result, disappear from actual speech).

Eckert (2008a:454) builds upon Silverstein's orders of indexicality to propose the idea of the *indexical field*, a "constellation of ideologically related [social] meanings, any one of which can be activated in the situated use of the variable." In Eckert's account, the social meanings of a given variable are not fixed but subject to ideological mediation in ways that can either draw upon pre-existing social meanings within the indexical field or extend the indexical field outward. The findings of the present study, by contrast, show that the CVS possesses core social meanings (sounding Californian, like a "Valley girl," and confident) that are not sensitive to context; by proposing the existence of empirically observable core social meanings that sit at the center of a given variable's indexical field, the present study adds a new layer to the concept of the indexical field (see 5.2.1. Social meaning, below).

Glossing Silverstein's *n* through (*n*+1)+1 levels as first- through third-order indexicality, Johnstone et al. show how variables originally used primarily by male, working-class speakers in Pittsburgh (e.g., MOUTH monophthongization, the *needs V-en* construction, the *you'uns* second-person plural) came to be associated with a notion of "Pittsburghese," a cultural object that implies some uniqueness to Pittsburgh ways of speaking. In early-1900s Pittsburgh, these features were unremarkable to their users because, as far as speakers were concerned, that was how "everyone said it" (88) within their dense, multiplex, locally based social networks. The first-order indexical facts of these variables' particular use by working-class males started to become noticed, and—mediated by an ideological connection between vernacularity and localness—this first-order indexicality gave rise to a second order wherein *you'uns*, etc., came to signal stylistically differentiable localness.

Geographic mobility, regional consciousness, and economic changes then provided the initial conditions necessary for third-order indexicality. As socially mobile Pittsburghers began to travel and meet individuals from outside of Western Pennsylvania, they found themselves "interact[ing] with people who sounded different and noticed how the Pittsburghers sounded" (93); moreover, the closure of steel mills in the late 20th century meant large numbers of Pittsburghers left the region for good, allowing for Pittsburgh-based social networks to interface (much more than previously) with individuals who could notice and comment upon Pittsburgh speech. Regardless of whether they stayed or left, third-generation Pittsburghers diverged from the primarily ethnic/religious self-conceptual schema used by their forebears, thinking of themselves instead in terms of class and regionality—developing, in other words, a shared imagined conception (Anderson 1983) of the working-class Pittsburgher. This third-order indexicality, the ability to describe what a Pittsburgher sounds like, meant that Pittsburghers

engaged in more talk about Pittsburgh talk, allowing for the creation of folk-linguistic artifacts such as websites on "Pittsburghese" (e.g., http://www.pittsburghese.com/). The end result is that, today, the youngest Pittsburghers may have their primary (or only) encounters with allegedly 'local' speech—that is, the variety that has been enregistered as Pittsburghese, even if not necessarily representative of the current first-order indexical facts-on-the-ground—not by interacting with speakers of Pittsburghese but by *learning of the existence* of Pittsburghese from folk-linguistic artifacts.

1.5.3. Enregisterment

Research in perceptual dialectology (see 1.4.2. Perceptual dialectology, above) reveals the existence of highly available varieties as social objects, as in "Southern talk" and "Jerseyites slang" seen in Figure 1.4 above. It is not a given that any particular regional dialect should become sufficiently visible to attract this sort of commentary; that is, the highly available nature of "Southern talk" (etc.) must be the outcome of a process of *enregisterment* by which a bundle of features becomes noticeable as a "socially recognized register of forms" (Agha 2003:231).

Johnstone et al.'s (2006) example of "Pittsburghese" in the previous section is actually a story of enregisterment, as *you'uns* and other variables traveled from first-order indexicality to a socially recognized package of linguistic practices symbolic of Pittsburgher identity. Although Johnstone et al. use an ordinal metaphor (*first*- through *third*-order indexicality), which implies an orderly progression of discrete indexical orders, it should be clear from the above description that the enregisterment of Pittsburghese was not an even or clean process. Just as Eckert (2000) shows that the Northern Cities Shift is not a monolith but the sum of individual agentive acts (see 1.5.4. Linguistic sense of place, below), enregisterment relies on the *circulation* of higher-order indexical associations. This is often accomplished via material *folk-linguistic artifacts* (Johnstone

2013, "linguistic artifacts" in the original); in Pittsburgh, the T-shirts that include lists of Pittsburghese lexical items (Johnstone 2009) constitute a paradigm example. Although Pittsburghese shirts are somewhat in the spirit of "folk dictionaries" (see e.g., Beal 2009:142–45), an older form of folk-linguistic artifact, a key feature is that they are wearable, allowing wearers to signal local identity. In order for Pittsburghese lexical items to even make it onto a shirt, however, the concept of Pittsburghese must be sufficiently entrenched that people will bother to pay money for a shirt representing it, which was only possible once Pittsburghese became a third-order indexical of local identity (163). Pittsburghese shirts thus exemplify the dynamism inherent to circulation—while they cannot come into being as consumer products without a certain level of Pittsburghese enregisterment, their production, consumption, and display further enregister Pittsburghese.

1.5.4. Linguistic sense of place

Linguistic features can index all sorts of social attributes (ethnicity, class, gender), but when studying indexical relationships between language and geography we can consider the central human-geographic notion of *place* within a sociolinguistic framework. In short, geographers differentiate *place* from *space* in terms of social meaning, in that any location can be a space, but only those imbued with some sense of human meaning are places:

> 40.46°N 73.58°W does not mean that much to most people. Some people with a sound knowledge of the globe may be able to tell you what this signifies but to most of us these are just numbers indicating a location—a site without meaning…Replacing a set of numbers with a name means that we begin to approach "place." If we heard that two planes had flown into 40.46°N 73.58°W it would not have quite the same impact as hearing that they had flown into New York, into Manhattan, into the Twin Towers.

> Cruise missiles are programmed with locations and spatial referents. If they could be programmed with "place" instead, with all the understanding that implies, they might decide to ditch in the desert. (Cresswell 2004:2)

Moving beyond first-order indexicality means that, as in Pittsburgh, speakers can utilize linguistic resources to index "being from Place X." In other words, within a certain community there is a notion of what it means to sound like you're from there—a *linguistic sense of place*. A linguistic sense of place includes not only linguistic properties of the variety (or varieties) associated with a certain place (lexicon, phonology, grammar) but also social evaluations of this variety (unmarked/standard vs. marked) and its speakers (class, intelligence, etc.). We can say that folk-linguistic artifacts such as Pittsburghese shirts are engaged in a process of *linguistic place-making*: the enregisterment of varieties that are primarily indexical of place.

A linguistic sense of place is never constructed in a vacuum, since, as in Pittsburgh, it is likely to have developed out of the foundations laid by other (ethnic, class, etc.) indexicality, and since indexicality at order n+1 is "always already immanent as a competing structure of values" (Silverstein 2003:194). In short, to sound like you're from Place X also calls up other social characteristics. Nowhere is this shown more clearly than in the pioneering work of Penelope Eckert (2000) in the high schools of suburban Detroit. Eckert's ethnographic approach revealed that social life in these schools was organized around a dichotomy between two locally salient communities of practice: *jocks* and *burnouts*. This dichotomy subsumed several oppositions: jock suburban-ness vs. burnout urban engagement, jock middle-class status vs. burnout working-class status, etc. The first opposition is relevant to the overall dialectological state of the Northern Cities Shift (then well underway in the area after having begun in the cities proper); whereas the older variables (e.g., LOT fronting) were well established in the larger metropolitan area, the

newer ones (e.g., DRESS backing) indexed urban affiliation. Among Eckert's speakers, the older variables showed patterns of stratification typical of changes from below (Labov 1990), with girls leading boys, but the newer variables were used more often by burnouts than jocks regardless of gender, as burnouts indexed their Detroit-oriented identities. But as burnout styles were defined only partially by geography (and only partially by language), these Detroit-based variables also took on (and built on) indexical associations such as working-class status, solidarity, etc. In general, then, Eckert's jocks and burnouts show how a linguistic sense of place may be embedded within complex indexical fields (Eckert 2008a). Central to this process is the inherently local nature of sociolinguistic variation; large macro-social (i.e., regional) patterns of language variation are not imposed on their users but are instead the product of their reproduction and recirculation in local settings through the indexical meanings they acquire (see e.g., Eckert & McConnell-Ginet 1992:470).

As with other types of folk linguistic awareness, a linguistic sense of place can vary in availability, ranging from the stylistic stratification typical of sociolinguistic markers (i.e., second-order indexicals) to the overt commentary and imitation that allows a place to show up on PD maps drawn by individuals who have never stepped foot in the community. In this sense, Pittsburgh's linguistic place-making is incomplete despite the considerably firm stereotype status of Pittsburghese within the city. For example, Preston's (1996b:305) composite maps of regions identified on Michigan and Indiana respondents' hand-drawn maps, which include regions that appeared on at least 15% of maps, do not include Pittsburgh; in other words, Pittsburgh simply does not register as a salient dialect region to perceivers merely two states away. The processes of circulation that have resulted in the enregisterment of Pittsburghese have thus failed (so far) to spread the idea of Pittsburghese to large portions of the US population.

1.6. Visions of California

Having previously reviewed the dialectological facts surrounding the English spoken in California (1.3. English language variation in California), it is now worth considering previous research—and the gaps left by previous research—on the social meanings surrounding California speech: indexicality (1.6.1), perceptual dialectology (1.6.2), and enregisterment (1.6.3).

1.6.1. California indexicality

Several studies on California Vowel Shift (or, rather, individual features thereof) have suggested a variety of indexical meanings for California vowels. Podesva (2011), for example, examines the speech of "Regan," a gay Asian American man from Southern California, in three different situations: a party, a dinner with a friend, and a professional meeting. Podesva analyzes several CVS variables (GOOSE fronting, GOAT fronting, TRAP backing, TRAP-N raising) and finds that Regan is most shifted in the party situation and most conservative in the meeting. Podesva's claim—which I revisit below (see 1.6.3. California enregisterment)—is that Regan's use of California vowels indexes a "partier" persona, which is connected to larger social meanings of Californians as being carefree and fun.

Although the CVS was (likely) originally an innovation by Caucasian speakers in California, Fought (1999) and Eckert (2008b) show that speakers of various ethnicities utilize (or resist) the CVS to index identities other than Whiteness (i.e., in a transformation of the indexical landscape). Fought found considerable variation in the use of GOOSE fronting by Chicano English speakers in Los Angeles. Discarding a purely ethnicity-based account of GOOSE fronting, Fought also finds social class to be unsatisfactorily explanatory, as some of the most advanced users of GOOSE fronting in her sample are low-income and some of the most conservative are middle-class. Instead, the social factor most explanatory of fronting is gang affiliation, with non-gang

individuals more likely to front and gang-affiliated individuals (including gang members) less likely; as with Eckert's (2000) jocks and burnouts (see 1.5.4. Linguistic sense of place, above), gang affiliation is a social dimension that is not represented in large sociological surveys but emerges as vitally important to this particular community.

Eckert (2008b) focuses on the TRAP nasal split among Chicana/o and Caucasian children at two Bay Area elementary schools. While the nasal split is (or was) a first-order indexical of ethnicity (with Chicana/o speakers resisting the split present amongst Caucasians), it has developed local second-order indexical values beyond ethnicity. At a predominantly Caucasian and middle-class elementary school, the nasal split indexes gender, being utilized more by girls, whereas at a poorer and more ethnically diverse elementary school, the nasal split is used only by children who are peripheral to the in-crowd. Both studies take place within a landscape where the CVS is the supraregional norm for Caucasian speakers, but in both studies speakers construct and redefine ethnicity-based indexicality via adoption of or resistance to the CVS.

Hall-Lew (2009) investigates competing linguistic senses of place in a single neighborhood: San Francisco's diverse Sunset District. Hall-Lew's ethnographic work uncovers two linguistic markets (Bourdieu 1977) in the Sunset: the Traditional Market, defined by the aging, Irish American residents who embrace a notion of San Franciscan distinctiveness from the rest of California (and affinity with the Northeastern US), and the Emergent Market, defined by a meshing of Caucasian and Asian cultures and an embedding of San Francisco identity within Northern California (103–09). Hall-Lew also finds resistance to two CVS features (the LOT/THOUGHT merger and back vowel fronting) among the Traditional Sunset Natives but acceptance among speakers who orient toward the Emergent Market. As a result, there is "a shift between the older and younger generations in terms of what forms index local authenticity"

(109). A more recent study on Hall-Lew's Sunset District data found that Chinese Americans have a smaller nasal split than Caucasians, which accords with the pattern in Eckert's predominantly Caucasian middle school (Cardoso et al. in press). These findings demonstrate that, with respect to a linguistic sense of place, spaces can acquire linguistic meaning at numerous scale levels, with speakers' neighborhood-based, city-based, and region-based identities often interacting with one another (see e.g., Becker 2009).

Podesva et al. (2015) found that Redding-area speakers' orientation to "town" vs. "country" (i.e., Redding vs. rural Shasta County) predicted the extent to which they used CVS variables; unexpectedly, country-oriented speakers led town-oriented speakers in fronting mid and back vowels, whereas town-oriented speakers led in the LOT/THOUGHT merger. Podesva et al. account for this result by positing that in Redding, the fronting of mid and high back vowels signals the speaker's orientation to the *Southern* Vowel Shift rather than the CVS, both of which share these features.[10] This interpretation suggests that apparent homogeneity in vowel production across California may mask heterogeneity in the social meanings of vocalic variation: "Each city has a different perceived character, and the emergence of this character as distinctive enables the speech styles of the city to become indexes of it" (180).

As discussed in the first paragraphs of this chapter, the production studies reviewed in this section leave open the question of whether listeners actually pick up on the social meanings that speakers evidently transmit in using CVS variables. For example, do Regan's friends react to his use of GOAT fronting by hearing him as a "partier"? If not, can we really claim that that is the

[10] The KIT-N/DRESS-N merger is unambiguously a Southern and non-CVS feature, but is found in parts of the Central Valley. As such, the authors could have tested their hypothesis that advanced speakers were orienting to the Southern Vowel Shift rather than the CVS by addressing the social patterning of the KIT-N/DRESS-N merger in Redding. This is an unfortunate omission.

social meaning indexed by Regan's GOAT fronting?[11] The present research empirically tests these production studies' claims about the social meanings of the CVS, finding that they are not supported by Californian listeners' evaluations of decontextualized stimuli (see 4.2.1. Guise, below). Rather than casting doubt on previous studies' claims, I propose in the Conclusion that the indexical field (1.6.1. California indexicality) provides a link between production and perception in the construction of social meaning (see 5.2.1. Social meaning, below).

1.6.2. California perceptual dialectology

Preston's (1996b) early PD research reveals a popular notion of uniquely Californian ways of speaking, as 17% of Michigan respondents identified a California speech region on hand-drawn maps (305).[12] (I do not claim that this 20 year old study stands in for the current state of affairs, but as I argue below, there is ample reason to believe that this popular notion of uniquely Californian ways of speaking has survived to the present day.) An inspection of six of Preston's individual maps (307–11) reveals interesting observations about outsiders' perceptions of California. On two maps, California enjoys no distinction from other states, grouped into (1) "Southwestern Neutral" on one and (2) "Damn Yankees" on another. On three maps, most or all of California forms its own region (with small parts of neighboring states presumably included more for ease of drawing than actual inclusion): (3) "not so much a diff[erent] lang[uage] but completly [sic] diff words: totally rad etc…," (4) "High Class partying Slobs & Stuck up

[11] Eckert (2008a:457) gives the example of two 1980s high schoolers in Palo Alto who partially aligned themselves with "New Wave" style by pegging their blue jeans, which they identified as indexing autonomy, but rejected the black eye makeup and clothing that they read as indexing adultness and sluttiness. Eckert does not address whether classmates of these girls apprehended these girls' stylistic moves, but I am skeptical that their intended social meanings were recognized by more than a handful of classmates. At a certain level, we must question whether social meaning is truly being constructed if stylistic moves are intelligible only to the speaker/wearer/transmitter.

[12] By comparison, 94% of Michigan respondents drew a South speech region, 54% drew a Northeast region, and 23% drew a Texas region.

sound,"[13] and (5) "Fer-Sher [*for sure*] Valley People." Another map circles a small area roughly west of Fresno: (6) "valley girl."[14] None of these labels employ an extraordinary level of metalinguistic detail; label 3 indicates an unavailability of phonological/prosodic features of California speech (in contrast to the same respondent's use of "whiney," "slow," and "nasal" elsewhere) but an overall evaluation of California as different somehow. Labels 4, 5, and 6 tap into the same carefree stereotypes indexed by Podesva's (2011) speaker Regan (see 1.6.1. California indexicality, above), with an added element of haughty affluence for label 4, and as circulated by the song "Valley Girl" (overtly for labels 5 and 6). In terms of correctness/pleasantness ratings, California ranks relatively high among respondents from Michigan and Indiana, but lower among Southerners (312–17).

A limited amount of PD research has investigated perceptions *by* Californians of language variation over space. Fought (2002) found that Californian students do not share Michiganders' notion of Upper Midwestern speech as "standard." In fact, it is questionable whether Californians share even the same underlying framework of "standardness;" whereas New England states were more likely to be labeled "proper English" (California received zero "proper English" labels), Western states (and California most of all) were more likely to be labeled "good/better/best English." Fought attributes this apparent contradiction to Californians' acknowledgment of marked speech within the state but overall linguistic security: "Californians seem to see their own speech in a fairly positive light, as natural and relaxed, but with its positive

[13] The map that includes label 4 is reproduced above as Figure 1.4 (see 1.4.2. Perceptual dialectology, above).
[14] As discussed below (1.6.3. California enregisterment), while the "Valley" in "Valley girl" originally referred to the suburban San Fernando Valley north of downtown Los Angeles, the geographical specificity of this stereotype was lost rather quickly, especially among non-Californian perceivers. The identification of western Fresno County as being the unique site of Valley girls is thus unfaithful to the original stereotype, but it is difficult to call it "inaccurate" per se. The metalinguistic label "Valley girl" itself probably helped to accelerate the loss of the stereotype's geographical specificity, as it does not specify the valley it refers to (and California has no short supply of valleys in all parts of the state).

value tempered by the idea that it is also not 'accurate' speech in some sense, and that it reflects the negative aspects of the surfers and the Valley girls" (132).

Bucholtz et al. (2007) restrict their analytic scope to California, collecting Californian students' hand-drawn maps of the state itself. The authors' qualitative analysis found four main types of category labels: geography, slang, languages, and social groups. With respect to geography, the state was most often cleaved into North vs. South, a distinction that typically surfaced as an opposition between the Bay Area and Los Angeles (338). Far Northern California and (especially) inland parts of the state, on the other hand, were subject to substantial erasure: "*Almost no one lives here*; *Does Anyone Live Out Here?*; *No man's land*; *Nothing- oppression, ennui, desert*; and *DEATH VALLEY (NO ONE SPEAKS)*" (338). This distinction also rested on the intensifier/quantifier *hella*, reported to be far more prevalent in the North and often derided by non-Northern Californians (343). Languages also surfaced as a labeling category, apparently to a much greater degree than in other map-drawing studies. The most salient labels were English and Spanish, the former more strongly associated with the North and the latter the South; Chinese and "Diverse" also featured prominently in the Bay Area, and few dialects of English received particular mention (other than "Ebonics" in Los Angeles and "Spanglish" inland and in San Diego) (340). With respect to social groups, the expected "surfers" and "Valley girls" were found in the South (the former also in the Central Coast), but most frequently mentioned was "hicks"—a Californian stereotype escaping the rest of the nation—in Northern California and inland. Other labels ("Mexicans" and "Latinos" in San Diego and inland, "laid-back" and "White" on the Central Coast, "hippies" in Northern California, "gangsters" in Los Angeles, "gays" in the Bay Area) received only a small share of the total. Moreover, while these social group labels may be interpreted as indicating awareness of recognizable varieties or features

associated with those groups (e.g., the label "Mexicans" indicating awareness of Chicano English), linguistic features received rather little overt comment, save for slang terms and languages. What thus emerges from these maps is a diverse linguistic sense of place for California, one in which "whiteness is unmarked and nonsalient while Mexicanness is marked and salient" (347) and social differences appear to be far more available than purely sociolinguistic ones.

D'Onofrio (2015) used a psycholinguistic task to assess whether listeners implicitly associated TRAP backing with the macro-social category label "Californian" or the persona label "Valley girl." In this task, listeners performed a lexical choice task with ambiguous LOT/TRAP words (e.g., *sock*/*sack*), which were created via resynthesizing vowels that were midway between /ɑ/ and /æ/. Listeners fixated on an icon between trials, and listeners' eye movements were tracked during trials. Listeners were in one of three conditions based on the social information conveyed via the inter-trial fixation icon: no social information (green vs. orange circle), macro-social (California vs. Michigan), and persona (Valley girl vs. nerd). The lexical choice results indicated an association between the Valley girl persona and TRAP backing, as the proportion of TRAP responses to ambiguous vowels was significantly greater in Valley girl trials. Both macro-social and persona information affected automatic perception, as listeners in both the Californian and Valley girl conditions spent significantly less time fixating on the LOT word than listeners in the baseline. Across conditions, listeners from the West chose the LOT word at a significantly greater rate, indicating that they were less attuned to the TRAP backing prevalent in California. (Condition and listener region did not significantly interact.) While this study fails to advance the argument that personae are more influential than macro-social categories in implicit

sociolinguistic perception, it provides evidence that listeners are aware of the association between TRAP backing and California.

1.6.3. California enregisterment

The results of Preston's (1996b) early PD work (described in the previous section) indicate a relatively well-established folk-linguistic notion among non-Californians that Californians speak differently than the rest of the US. It is apparent that several notable cultural artifacts have served to establish and/or circulate this notion, the first of which seems to have been Frank Zappa's popular 1982 song "Valley Girl" (the "Valley" referring to the suburban San Fernando Valley north of downtown Los Angeles). According to an Associated Press wire story (1982), Zappa's daughter Moon based the spoken-word portions of the song on "having met dozens and dozens of Valley girls and absorb[ing] everything that [she]'d overheard from parties and bar mitzvahs and the Galleria [a shopping mall]." While Moon characterized her portrayal as "Valley lingo,", and while her portrayal does present several unique slang terms/phrases ("tubular," "for sure," "gag me with a spoon"), it clearly extends beyond the lexicon; the character also encompasses unique features of prosody (very high pitch maxima), phonology (GOAT fronting, centralization of FLEECE-L in *really*), and discourse (focuser and quotative *like*). Moreover, Moon's spoken-word monologues convey a carefree, aloof privilege via their content ("Like my mother like makes me do the dishes, it's like so gross…it's like somebody else's food, y'know, it's like grody, grody to the max!"). In other words, the song offers listeners a place-linked speech register, prepackaged with an image of its typical speakers and how listeners should feel about them—not to mention a convenient label by which this stereotype could circulate.

Since 1982, prominent media depictions have reinforced a popular connection between (Caucasian) Californians and certain ways of speaking. These ways of speaking include a "surfer" stereotype advanced by the 1982 movie *Fast Times at Ridgemont High* (embodied by Sean Penn's long-haired Jeff Spicoli)[15] and a "slacker" stereotype advanced by the 1989 movie *Bill and Ted's Excellent Adventure*. More recently, Saturday Night Live has featured a recurring sketch "The Californians," a mock soap opera in which the CVS-tinged (see fn. 17) dialogue between affluent Caucasian (bleach-blonde) Southern Californians consists primarily of the combinations of freeways and surface roads required to reach one point from another (thus adding a topical dimension to "Californian" speech).[16]

Notably, all four of these portrayals, from "Valley girl" to "The Californians" center on *Southern* California, through references to both geography (e.g., specific freeways) and practices perceived to typify the Southern California experience (e.g., surfing, consumerism, consumption of local produce). Whether intended or not, these linguistic stereotypes have often been consumed and circulated by non-Californians as representing California speech generally (with Southern California marked as the exemplar of the Californian lifestyle and experience), as suggested by Preston's (1996b) hand-drawn PD maps showing Valley girls in the entire state (see 1.6.2. California perceptual dialectology, above). For Moon Zappa, this geographical erasure was unintentional, as she described her "Valley girl" portrayal as representing her friends' "inside joke about people who live in Encino" (Associated Press 1982), a San Fernando

[15] Although the song "Valley girl" (released in June 1982) entered the public consciousness before *Fast Times* (released in August 1982), the lag between movie filming and release means that Sean Penn's acting must have taken place prior to the song's release. In other words, the Spicoli character did not build on the Southern California stereotypes popularized by the song, meaning that these portrayals must have been developed independently of one another.

[16] According to the Saturday Night Live Wiki (http://snl.wikia.com/wiki/The_Californians), the sketch first aired April 14, 2012, with five more iterations between May 2012 and May 2013, making it one of the most frequently recurring sketches in recent seasons.

Valley neighborhood of Los Angeles. The song's quick nationwide popularity (which came as a surprise to Moon) was not due to a widespread awareness of inter-neighborhood distinctions and stereotypes within Los Angeles, of course, but because "Valley girls" were interpreted as an adequate substitute for Southern Californians, who were then interpreted as an adequate substitute for Californians generally. (This geographical erasure appears to be intentional for Saturday Night Live, as the sketches, which clearly take place in Southern California, were titled "The Californians" rather than "The Southern Californians.")

Podesva (2011) argues that his speaker Regan indirectly calls up these character types in shifting toward California vowels (see 1.6.1. California indexicality, above); that is, Regan is not voicing Jeff Spicoli per se, but the fun and carefreeness that speakers like Spicoli represent (not to mention "Spicoli" as a popular movie character and thus a cultural artifact that contributed to this linguistic place-making in the first place). But this account is incomplete for at least two reasons. First, folk-linguistic artifacts like Zappa's original Valley Girl and Spicoli can vary along lines of accuracy, just as folk-linguistic awareness can. It is not clear whether Southern California English featured TRAP backing and TRAP-N raising in 1982, but neither character features either of these variables in their speech;[17] Regan, however, uses both in constructing his 'partier' persona. Second, Podesva compares the enregisterment of California speech to Pittsburghese, but these processes have had rather different outcomes (so far). As mentioned above, PD map-drawing studies provide evidence that California enjoys a more firmly

[17] In fact, both Spicoli and Moon Zappa's "Valley girl" variably show some *backing* of TRAP-N, raising the possibility that this was a short-lived change in Southern California. Pratt and D'Onofrio (2014) found that Kristen Wiig's character in "The Californians" (which began airing after Podesva's analysis was published) also backs both TRAP and TRAP-N (whereas Fred Armisen's character *raises* both TRAP and TRAP-N). Other vowels used by "The Californians" also suggest overgeneralizations of other CVS features (e.g., FLEECE/FACE backing as an overgeneralization of the movement of front lax vowels). More research needs to be done on the relationship of popular folk-linguistic portrayals to changes in progress.

established linguistic sense of place among outsiders than does Pittsburgh. This state of affairs is partially due to the nature of the producers and consumers of folk-linguistic artifacts (see e.g., Beal 2009). In the case of Pittsburgh, small-scale vendors sell shirts to individuals, who have a relatively limited audience when wearing their shirts (as constrained by the limited number of individuals who can be interacted with, the limited number of settings in which a T-shirt may be appropriately worn, etc.); conversely, movie and television characters like Spicoli are produced for consumption by wide audiences. As a result, it is not clear that "Valspeak" or "Surfer talk" is available to speakers as third-order indexes of identity the same way that Pittsburghese is. Rather than claiming to speak these varieties as a matter of local pride, speakers are more likely to claim that they are somebody else's.[18,19]

Folk-linguistic artifacts produced by and for Californians, on the other hand, tend not to portray Californians as Valley girls or surfer dudes. A 1984 *San Francisco Chronicle* article details "a distinctive San Francisco way of talking" that is in danger of becoming "as dead as Latin" (Nolte 1984). Unlike accounts from Bucholtz et al.'s map-drawing results, the author presents a folk-linguistic discourse in which San Francisco itself is singled out rather than a larger Bay Area/Northern California speech area. According to the article, authentic San Franciscan speech is a combination of phonological reduction and a familiarity with local human

[18] A short Associated Press item ("Help for Valspeakers") running the same day as the Moon Zappa piece details the reparative efforts of a Beverly Hills speech pathologist who begins her therapeutic sessions by recording "patients" and playing back the recordings for them: "'A lot of them are shocked when they see themselves on tape,' she said. That is followed by a battery of audio and video lessons to help the patient relearn standardized English." The shock that these "patients" experience is the realization that somewhere along the way they lost their grip on "standardized English" and became "Valspeakers." This anecdote also points to the CVS's status as a change from below.

[19] Contrary to this general observation is the recent status of *hella* as a linguistic feature that is claimed by its users as a conscious identity marker, as in the song "I Say Hella" by Oakland-born rapper Bobby Brackins: http://youtu.be/OVRuKvT9naY. In a parallel to Pittsburgh, several vendors now sell shirts and hooded sweatshirts reading "I hella ♥ California/Cali/CA/Oakland/The Bay/San Francisco/SF" (e.g., https://www.spreadshirt.com/i-hella-love-san-francisco-hoodie-A6794850), a variation on the "I ♥ NY" meme. These developments are a far cry from California speakers claiming to sound like Valley girls, though.

geography: "The largest city in Santa Clara County is 'Sannazay,' not 'San Jose.' Sannazay is near Sannacruise. To get there, you have to go Down the Peninsula, past South City, Sammateo, Rewoodcity and a whole buncha other towns." Of course, phonological reduction not only is characteristic of many vernaculars (Kroch 1978) but also features prominently in folk-linguistic depictions of other American cities.[20] What appears to be at work in the Nolte article is the same discourse of local-authenticity-via-vernacularity also found in folk-linguistic artifacts from the very cities against which the article asserts an oppositional identity. These folk beliefs about local San Francisco speech are by no means Nolte's alone, however; as Hall-Lew (2009:63–67) found, native Sunset residents over 40 recalled a San Francisco accent—lacking in detail—associated with Irish Catholics and with its epicenter in the Mission District.[21] Younger Sunset natives, on the other hand, felt that San Francisco was either linguistically unremarkable or placed it in opposition to Southern California by the latter's use of *the* before freeway numbers (Geyer 2001) and nonuse of *hella*.

To summarize, California appears to inspire conflicting visions as a linguistic sense of place. To outsiders, it is a land of youthfully vacuous speakers, mostly from Southern California, with fronted GOOSE and GOAT. To insiders, it is characterized primarily by its populated areas, linguistically diverse, and not quite proper but not terribly marked in general, either. The enregisterment of California speech thus remains incomplete, as for many residents of the state "California English" appears not to be a social reality.

[20] e.g., "Phillyspeak": http://archives.citypaper.net/articles/081497/article008.shtml
[21] Hall-Lew noted that Nolte's article was still in circulation as a folk-linguistic artifact among San Franciscans at the time of her fieldwork, more than 20 years after the article's original publication. The article's staying power as a folk-linguistic artifact is a testament to its popularity as a nostalgic description of 'authentic' San Franciscan vernacularity.

1.7. The present research

1.7.1. Research questions

The present research seeks to better understand how Californian listeners participate in the construction of the social meanings of the Californian Vowel Shift, especially the CVS's status as a change from below and the salient folk-linguistic profile of California stereotypes. While this basic question suggests numerous possible directions, the present study focuses on the following research questions:

1. To what extent (if any) do Californians perceive the California Vowel Shift as Californian?
2. What evaluations do Californians attach to the California Vowel Shift? How do these evaluations compare to popular portrayals of California speakers?
3. To what extent (if any) do listeners from different regions of California evaluate speakers of California English differently?

1.7.2. Outline of the dissertation

Chapter 2 describes a pilot study that I conducted in summer 2013 to assess the viability of a dialect recognition task for approaching the above questions (Villarreal in press). After discussing the design, results, and implications of the pilot, I outline areas in which the pilot's basic design would later be expanded and/or improved to allow the main dissertation study to better address the research questions.

Chapter 3 describes the methods of the main study (conducted from spring 2014 to spring 2015) in three stages: sociolinguistic interviews, a pretesting task, and the main experimental task. Both the interviews and pretesting stages facilitated the main task, and I briefly describe the

results of each. I then discuss the methodological details of the main task, including the procedure by which excerpts were acoustically manipulated to create matched-guise stimuli, as well as my analytic methods.

Chapter 4 describes the results of the main experimental task. I outline how guise and other predictors such as speaker ethnicity and listener region affected perceptions of speakers' regional origin and personal characteristics. I also analyze the intersection of these two types of responses.

Chapter 5 concludes the dissertation by discussing how the main task's results bear on the original research questions. I also suggest ways that this task, and tasks like it, contribute to theory and methods in the arenas of social meaning, perceptual dialectology, and language change. I end by outlining future research directions with this corpus of data and beyond.

Chapter 2
Pilot study

The goal of the pilot study was to assess the viability of using a dialect recognition task (modeled on that of Williams et al. 1999) for studying the perceptual dialectology of California English.[1] In this task, listeners heard samples of speakers from different regions of the state, guessed where speakers were from (regional identification, and rated speakers on affective scales (attitudes). The data were analyzed by first looking at regional identification and attitudes separately, then investigating the interaction of these two types of responses. The results demonstrated that a dialect recognition task was indeed suited to revealing interesting trends in the perceptual dialectology of California English; the speech of the San Francisco Bay Area was perceived to have greater correctness and Southern California greater pleasantness, and listeners from both of these regions believed their own region to be central to a notion of California speech.

[1] A version of this chapter is in press as "'Do I sound like a Valley Girl to you?' Perceptual dialectology and language attitudes in California," chapter 5 of *Speech in the West*, vol. 1, ed. by Valerie Fridland, Betsy Evans, Tyler Kendall, and Alicia Wassink (Durham, NC: Duke University Press).

This chapter provides an overview of the pilot study, as the main experimental task was modeled on the pilot task, with substantial improvements. In the following sections, I describe the methods (2.1) and results (2.2) of the pilot, briefly consider how the pilot connects with previous research (2.3), and discuss methodological gaps revealed by the pilot and improved upon in the main study (2.4).

2.1. Methods

Using Williams et al.'s (1999) dialect recognition task as a model, the pilot study asked respondents to 'place' speakers in one of five regions (four Californian regions and one catchall "outside California" region) and rate speakers on 14 semantic differential scales for affective traits. Unlike Williams et al.'s task, the stimuli in the pilot study were identical in content, meaning listeners responded primarily to vocalic variation among speakers.

The following sections describe the pilot study's methods: the regions of California from which stimuli were drawn (2.1.1), the recording and preparation of stimuli (2.1.2), the design of the recognition task itself (2.1.3), and the analytic methods (2.1.4).

2.1.1. Regions

The four Californian regions defined for the purposes of the pilot study were Far Northern California, the (San Francisco) Bay Area, the Lower Central Valley, and Southern California; Figure 2.1 displays these regions.[2] Although the existing evidence for intrastate differences in California English is minimal (see 1.3.1. Intrastate variation in California English, above), these regions were selected on the basis of economic, geographic, and ethnolinguistic

[2] Note that these regions exclude some parts of the state, mostly sparsely populated areas (the Sierra Nevada mountains, Mojave Desert), as well as Greater Sacramento, a region that could arguably be called part of the Central Valley, an extension of the Bay Area, or its own region.

differences—in other words, sociocultural factors that are often responsible for dialect differentiation (Wolfram & Schilling-Estes 2006:29–35)—as well as their differential folk-linguistic status.

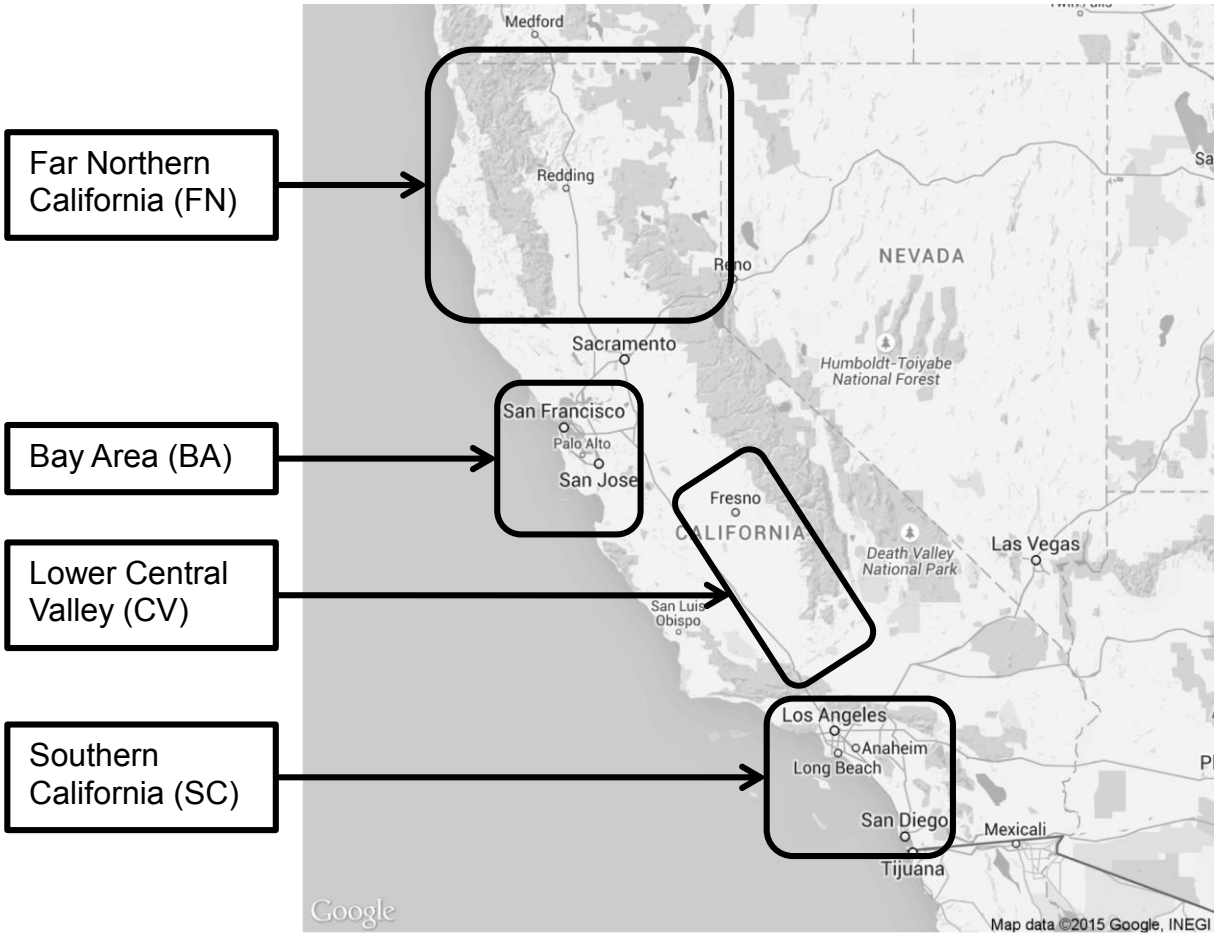


Figure 2.1. Californian regions and abbreviations in the pilot study.

Perhaps the most drastic differences between these regions are in population (displayed in Table 2.1) and urbanization. In terms of population, Southern California, which contains the US's 2nd, 13th, and 17th most populous metropolitan areas (Los Angeles, Riverside, and San Diego), dwarfs all the other regions. The Bay Area, anchored by San Francisco and Oakland (the US's 11th most populous metro area), is also heavily urbanized; like Southern California, most of its non-mountainous land is urban/suburban development. The Lower Central Valley contains

two cities, Fresno and Bakersfield, whose combined metro population is nearly 2 million, but otherwise the region is mostly rural farmland. Whereas Far Northern California's population is greater than eight US states, it is spread over the greatest land area, much of which is farmland or (especially) undeveloped forest. While California has modern transportation and telecommunications infrastructure, the large size of the state, as well as physical features such as the Diablo Range, serve to limit persistent, prolonged contact between speakers from different regions.

Region	Population	Land area	Density
Far Northern California	1,048,632	42,544	24.65
Bay Area	9,038,767	10,132	849.52
Lower Central Valley	2,603,578	22,441	116.02
Southern California	21,813,719	38,168	571.52

Table 2.1. California regions by population, land area (square miles), and density (persons per square mile).
Land area is in square miles; density is in persons per square mile. Population data from US Census Bureau July 1, 2014 estimates (http://www.census.gov/popest/data/counties/totals/2014/CO-EST2014-01.html). Land size data from US Census Bureau State and County QuickFacts (http://quickfacts.census.gov/qfd/).

Settlement and language contact (both past and present) also vary between these regions. Widespread settlement in California by English speakers did not begin until the Gold Rush of the mid-19th century, which brought settlers into the Bay Area and Far Northern California. San Francisco was subsequently settled by migrants from the eastern US and all over the world, with 20,000 Chinese immigrants arriving by 1852 alone (Rolle 1987:166–69). Southern California has experienced several population booms, first attracting Easterners and Midwesterners with its favorable climate in the late 1800s (Rolle 1987:275–83), then enjoying the rise of Los Angeles harbor as a crucial port (362), an increased naval presence in the 1940s, and booming postwar industry (447). In the 1930s, the Lower Central Valley experienced an influx of Oklahoma and

Arkansas farmers displaced by the Dust Bowl.[3] In more recent years, California has become increasingly diversified thanks to increased international immigration and domestic out-migration (Gray & Scardamalia 2012). As a result, speakers in different regions experience different language contact situations, as nearly half of all residents of the Bay Area, Lower Central Valley, and Southern California reside in households where English is not spoken (Table 2.2).

Region	English	Spanish	Other
FN	86.2%	10.2%	Hmong (0.5%), German (0.3%), Chinese (0.3%)
BA	58.9%	18.5%	Chinese (6.4%), Tagalog (3.4%), Vietnamese (2.2%)
CV	55.5%	37.8%	Hmong (1.1%), Other Indic Lgs (0.9%), Tagalog (0.9%)
SC	51.8%	33.8%	Chinese (2.6%), Tagalog (2.2%), Vietnamese (1.6%)

Table 2.2. Percentage of English, Spanish, and other languages spoken at home by region.
Region abbreviations as in Figure 2.1. "Other" column lists top three languages other than English and Spanish. Data from American Community Survey (US Census Bureau) 2010–2014 five-year estimates, table B16001, available at http://factfinder2.census.gov/.

In addition to differences in sociocultural factors, these regions enjoy different folk-linguistic status among both Californians and non-Californians (see 1.6.2. California perceptual dialectology, 1.6.3. California enregisterment, above). Californians typically construct difference along a north vs. south dichotomy (Bucholtz et al. 2008; Fought 2002), often erasing the less-urbanized Far Northern California and Lower Central Valley (Bucholtz et al. 2007), whereas to outsiders Southern California linguistic stereotypes such as the "Valley girl" tend to stand in for the state as a whole (e.g., Preston 1996b).

[3] It is likely that this migration pattern is responsible for the presence of the KIT-N/DRESS-N merger in Fresno and Bakersfield (see 1.3.1. Intrastate variation in California English, above), and both facts are likely related to Californians' identification of Lower Central Valley speakers as "hicks" (see 1.6.2. California perceptual dialectology, above).

2.1.2. Speakers and stimuli

Stimuli were short sound clips drawn from 12 Caucasian speakers in two corpora: the Nationwide Speech Project corpus (NSP; Clopper & Pisoni 2006) and a small corpus of newly recorded speakers. Two speakers (one female and one male) represented each of the four Californian regions, and four speakers (two females and two males) represented non-Californians. Seven of the 12 speakers were recorded as part of the NSP corpus: the male Bay Area speaker, both Southern California speakers, and all four non-Californian speakers. In order to complete the stimulus sample, ten additional Californian speakers, students at UC Davis (UCD), were recruited via social networks. Speakers were required to fulfill several criteria: self-identifying as Caucasian, speaking English natively, having lived in the same region their entire lives (prior to attending UCD), being between 18–29 years old, and reporting normal speech and hearing. Five of these UCD speakers were chosen for the recognition task based on their regional origin and use of Californian/Western vocalic features. Four speakers from the NSP's Mid-Atlantic region were chosen for the "outside California" region based on a lack of Californian/Western vocalic features. Table 2.3 displays information on the speakers.

Speaker	Region	Corpus	Gender	Age	Hometown	Trial
FN1	CA	UCD	F	21	Red Bluff	1
FN2	CA	UCD	M	28	Grass Valley	7
BA1	CA	UCD	F	21	San Geronimo	6
BA2	CA	NSP	M	19	San Mateo	3
CV1	CA	UCD	F	22	Bakersfield	9
CV2	CA	UCD	M	21	Fresno	8
SC1	CA	NSP	F	20	Los Angeles	4
SC2	CA	NSP	M	18	Santa Clarita	11
Out1	MA	NSP	F	18	Middletown, NJ	5
Out2	CA	NSP	F	18	Oradell, NJ	12
Out3	CA	NSP	M	18	Long Island, NY	10
Out4	CA	NSP	M	18	Plainview, NY	2

Table 2.3. Speakers included in pilot recognition task.
In "Region" column, CA indicates California, MA indicates Mid-Atlantic. "Trial" is order of presentation in recognition task.

The UCD speakers were recorded following similar procedures to the NSP, with the author overseeing all recording sessions. Unlike the NSP, UCD speakers only participated in one task: recording the 102 sentences from the NSP's high-probability sentence list in randomized order. Table 2.4 gives the four sentences that were chosen for the recognition task stimuli (in the order in which they were presented) and the 10 CVS/Western features represented across these four sentences. For each speaker (NSP and UCD), these sentences were spliced out of the speaker's original sound file and spliced into a stimulus sound file with 1 second inter-sentence intervals. As there was variation across speakers in baseline speech intensity and background noise, Praat (Boersma & Weenink 2015) was used to equalize files' intensity level to 60 dB and equalize background noise. All stimuli were around 10 seconds in duration.

Sentence	NSP #	Features
Follow this road around the bend.	9	GOAT, DRESS, DRESS-N
The bloodhound followed the trail.	43	STRUT, FACE-L
The swimmer's leg got a bad cramp.	83	KIT, DRESS, DRESS-G, TRAP, TRAP-N
Ruth had a necklace of glass beads.	36	GOOSE, TRAP, DRESS, TRAP

Table 2.4. Sentences chosen for stimuli and CVS/Western features represented in each.
"NSP #" is the sentence number in the NSP. Underlines indicate location of features, double underlines represent possible overlaps (e.g., the vowel in *bend* could undergo either DRESS lowering or the KIT-N/DRESS-N merger).

In controlling for the content of the sound clips, the pilot study differed from Williams et al. (1999), which used excerpts from speakers' spontaneous narratives for stimuli. Listeners' responses were thus tied only to phonological/prosodic differences among speakers, and the primary differences between speakers were vowel features. For reasons discussed below (see 2.4. Methodological changes for main study), the main study did not use stimuli with identical content, but instead controlled for topical variation (while allowing for minimal variation in content) by drawing stimuli from responses to a cartoon retell task (see 3.2.1. Retell task design, below).

In order to quantify these vowel differences, I coded each stimulus impressionistically for the presence or absence of the CVS/Western features in each sentence (Table 2.4) using a three-point scale: 0 for non-shifted, 0.5 for partially shifted, and 1 for shifted; for DRESS lowering, for example, [ɛ] would be coded 0, [ɛ̝] would be coded 0.5, and [æ̝] or [æ] would be coded 1. All of the non-Californian speakers also raised the vowels in *bad* and *glass* to varying degrees, a feature characteristic of a Mid-Atlantic short-a system (Labov et al. 2006b), so this feature was also coded for. These codes were then converted to indices to determine the degree to which stimuli contained CVS features (divided into front and back vowels), non-CVS Western features, and Mid-Atlantic *bad* raising. Indices were computed by dividing the sum of codes by the number of tokens and multiplying by 100; for example, CV1 was coded 1 for GOOSE, 0 for GOAT, and 0.5

for STRUT, so her CVS back index was 50. The CVS index was then computed by averaging the CVS front index and CVS back index. Speakers' indices are displayed in Table 2.5.

Speaker	CVS front	CVS back	CVS	Other West	*bad*
FN1	25	67	**46**	33	0
FN2	25	33	**29**	0	0
BA1	81	50	**66**	0	0
BA2	25	0	**13**	17	0
CV1	25	50	**38**	100	0
CV2	31	83	**57**	0	0
SC1	31	100	**66**	0	0
SC2	19	33	**26**	17	0
Out1	13	0	**6**	0	100
Out2	19	0	**9**	17	75
Out3	0	17	**8**	17	50
Out4	6	17	**11**	0	75

Table 2.5. Indices by speaker.

Speakers BA1 and SC1, for example, were the heaviest user of California-shifted vowels, with CVS indices of 66, whereas BA2 was the least California-shifted Californian, with a CVS index of 13. The CVS and *bad* indices clearly differentiate the Californian and non-Californian speakers, all of the non-Californian speakers had lower CVS indices than the least California-shifted Californian, while all of the non-Californians but no Californians had at least some *bad* raising. CV1 had a high rate of non-CVS Western features and was the only speaker to exhibit the KIT-N/DRESS-N merger; a second-generation Bakersfieldian, CV1 patterns in this respect as ANAE predicts (Labov et al. 2006b:68). Aside from CV1's KIT-N/DRESS-N merger, however, it is difficult to determine whether the Californian speakers are exemplars of their respective regions within the state due to the relative lack of information on intrastate vocalic differences. However, this research attempts to ascertain whether Californians attach social meaning to vocalic features in the same way that Bucholtz et al. (2007) found that they attach social meaning to lexical differences.

2.1.3. Recognition task design

Fifty-three UCD undergraduates participated in the recognition task. The task consisted of 12 trials (following a practice trial) in which listeners guessed where each speaker was from and rated speakers from 1–6 on 14 semantic differential scales (Table 2.6). Most scales were selected to correspond to the "correctness" and "pleasantness" groupings found in PD studies (Preston 1996b). Each trial lasted 60 sec, with the sound clips played twice. Stimuli were presented in the same order to all listeners (displayed in the "Trial" column of Table 2.3). Listeners then filled out a brief questionnaire, which included a space for listeners to list where they had lived and which of those places they considered their hometown.

Category	**Scales**	
Correctness	*smart–dumb*	*good English–bad English*
	educated–uneducated	*fast–slow*
	confident–not confident	*unaccented–accented*
Pleasantness	*attractive–unattractive*	*polite–rude*
	friendly–unfriendly	*likeable–not likeable*
Self-similarity	*speaks like me–doesn't speak like me*	*familiar–unfamiliar*
Place	*Californian–not Californian*	*rural–urban*

Table 2.6. Semantic differential scales and categories used in perceptual task.

2.1.4. Analysis

Listeners' questionnaire responses were coded for hometown and mobility. Hometown categories were Far Northern California, the Bay Area, Lower Central Valley, Southern California, other California locations, non-Californian US locations, and international locations. Mobility categories were non-mobile (listeners had never lived outside their home region of California), mobile across multiple regions of California, mobile across multiple states, and mobile across multiple countries.

The data analysis fell into three broad categories: regional identification, affective scale ratings, and the interaction of regional identification and attitudes. Accurate-recognition rates were calculated for individual speakers, speaker regions, listener regions, and listener mobility. Identification rates were also investigated independently of accuracy, with special attention paid to *perceived region* (the region into which a plurality of listeners placed a speaker),[4] *ingroup identification rates* (the proportion of listeners from a given region who identify a speaker as being from their home region), and *outside identification rates* (the proportion of California listeners who identify a speaker as being from outside the state).

Average ratings for affective traits were calculated by region in two ways: speakers' actual region and speakers' perceived region. To assess the presence of *claiming* effects in California (see 1.4.3. Dialect recognition, above), average ratings were also calculated for listeners from the two regions most heavily represented in the listener sample (the Bay Area and Southern California), and correlation tests were performed to compare average ratings with ingroup identification rates. Finally, correlation tests were performed to compare Californian listeners' average ratings with outside identification rates.

2.2. Results

A one-way ANOVA analysis of recognition revealed a significant effect of listener mobility on accurate recognition, $F(3, 49) = 3.83$, $p < .05$. Post-hoc Scheffé tests showed that listeners who had lived outside the US were significantly less accurate than those who had moved among Californian regions and those who had lived in other US states ($ps < .05$). As a result, these 12 listeners were removed from the sample for the analyses described in this section;

[4] In the main study, trials were categorized into perceived regions based on the regional identification response for each trial.

after removing these listeners, there was no significant effect of listener hometown, $F(4, 35) = 1.01$.

In the remaining 41-listener sample, Far Northern California (n = 1), the Lower Central Valley (2), and outside California (4) were underrepresented compared to the Bay Area (11), Southern California (13), and Greater Sacramento (10). As a result, it was feasible to analyze ingroup identification only for the Bay Area and Southern California.

2.2.1. Regional identification

The overall rate at which speakers' regions were accurately recognized was 31.2%, consistent with Clopper and Pisoni's (2004) 30% and Williams et al.'s (1999) 25%. As Table 2.7 shows, few listeners accurately recognized Californian speakers' region; for example, 15.8% of listeners identified CV2, the male Lower Central Valley speaker, as from the Lower Central Valley. For six of eight Californian speakers, their perceived region (the region into which a plurality of listeners placed a speaker) did not match their actual region. For example, a plurality of listeners placed FN1, the female Far Northern California speaker, in Southern California (39.0%), whereas 19.5% of listeners placed FN1 in Far Northern California. In contrast, listeners accurately recognized three of four non Californian speakers; more than two-thirds of listeners identified Out3 as from outside California. Accurate-recognition rates ranged from 9.8% (BA1) to 70.0% (Out3).

Speaker	Actual region		Perceived region	
FN1	Far North	19.5%	So. Cal	39.0%
FN2	Far North	19.5%	Bay Area	26.8%
BA1	Bay Area	9.8%	Far North	34.1%
BA2	Bay Area	22.0%	Far North	31.7%
CV1	Cent. Valley	27.5%	Cent. Valley	27.5%
CV2	Cent. Valley	15.8%	Far North	42.1%
SC1	So. Cal	48.8%	So. Cal	48.8%
SC2	So. Cal	24.4%	Bay Area	34.1%
Out1	Outside	61.0%	Outside	61.0%
Out2	Outside	25.0%	Cent. Valley	30.0%
Out3	Outside	70.0%	Outside	70.0%
Out4	Outside	31.7%	Outside	31.7%

Table 2.7. Percentage of listeners who placed speakers into their actual region and perceived region.
Speakers whose perceived region matched their actual region highlighted in gray.

In order to test whether listeners associated certain regions with the CVS or other features, Pearson correlations were calculated between the indices in Table 2.5 and the rate at which speakers were placed into each region. There was a significant positive correlation between speakers' CVS index and the percentage of listeners who identified them as being from Southern California (r = .69, one-tailed t_{10} = 3.05, p < .01), although there was no significant correlation for the CVS front and CVS back indices; in other words, speakers who exhibited more CVS features were identified as being from Southern California at a higher rate.

Table 2.8 gives rates of ingroup identification by Bay Area and Southern California listeners, as well as outside identification by all Californian listeners. Notably, over three-fourths of listeners from Southern California identified SC1 as a Southern Californian (76.9%), whereas no listeners from the Bay Area identified BA1 as from the Bay Area. Listeners from both the Bay Area (45.5%) and Southern California (53.8%) identified FN1 as an ingroup speaker. It is not clear how these ingroup identification patterns relate to vocalic variation within the sample;

while Bay Area ingroup identification correlated negatively with CVS indices and Southern California ingroup identification correlated positively with CVS indices, these correlations did not reach significance.

Not surprisingly, Californian listeners identified the non-Californian speakers as from outside California at a higher rate than they identified Californian speakers as from outside the state; for example, 72.2% of Californian listeners identified Out3 as from outside California, compared to just 2.7% for FN1 and SC1. The Californian speaker with the highest outside identification rate was CV2 (25.7%), which surpassed that of Out2 (19.4%). Outside identification rates correlated significantly with *bad* raising indices ($r = .72$, one-tailed $t_{10} = 3.30$, $p < .005$); in other words, listeners were generally more likely to identify speakers with *bad* raising as being from outside California.

Speaker	Ingroup identification: Bay Area	Ingroup identification: So. Cal	Outside identification
FN1	45.5%	53.8%	2.7%
FN2	9.1%	7.7%	16.2%
BA1	0%	30.8%	8.1%
BA2	36.4%	23.1%	8.1%
CV1	27.3%	30.8%	8.1%
CV2	0%	8.3%	25.7%
SC1	18.2%	76.9%	2.7%
SC2	45.5%	30.8%	0%
Out1	9.1%	15.4%	59.5%
Out2	27.3%	23.1%	19.4%
Out3	9.1%	0%	72.2%
Out4	9.1%	15.4%	32.4%

Table 2.8. Ingroup identification rates for Bay Area and Southern California listeners and outside identification rates for all Californian listeners.
Cells for which speakers were rated by listeners from the same home region highlighted in gray.

2.2.2. Language attitudes

Tables 2.9 and 2.10 give average affective scale ratings for each region; whereas Table 2.9 shows average ratings based on speakers' *actual* region (where speakers were actually from), Table 2.10 shows average ratings based on speakers' *perceived* region (perceived regions listed in Table 2.7). (Recall that scales ranged from 1–6.)

Turning first to Table 2.9, it is clear that Southern California speakers enjoyed favorable attitudes across the board, ranking highest on correctness items, pleasantness items, self-similarity items, and *Californian*. Because non-Californians ranked lowest on almost all scales, it is useful to examine the second-lowest average ratings for each scale. Here, it appears that the Lower Central Valley is the least prestigious Californian region (especially on pleasantness) and that the Bay Area also lacks prestige. The scale *rural–urban* (neither side of which was assumed *a priori* to correspond to correctness, pleasantness, or 'Californian-ness') shows the opposite

pattern of the other scales, with Southern California ranking lowest and non-Californians ranking highest on rurality; it thus appears that urbanness carries an element of linguistic prestige to these listeners.

		Actual region				
	Scales	**Far North**	**Bay Area**	**Cent. Valley**	**So. Cal**	**Outside**
Correctness	*smart*	3.94	3.66	3.72	4.31	3.58
	educated	4.35	3.96	3.96	4.54	3.65
	confident	3.85	3.46	3.61	4.23	3.46
	good English	4.59	4.15	4.08	4.65	3.89
	fast	3.48	3.52	3.38	3.83	2.76
	unaccented	3.96	3.97	3.81	4.35	3.13
Pleasantness	*attractive*	3.90	3.44	3.12	4.16	3.12
	friendly	4.33	3.64	3.60	4.57	3.62
	polite	4.43	3.77	3.68	4.52	3.87
	likeable	4.21	3.59	3.42	4.34	3.62
Self-similarity	*speaks like me*	3.47	3.14	2.82	3.73	2.34
	familiar	4.04	3.91	3.69	4.49	3.20
Place	*Californian*	4.44	4.53	3.83	4.99	2.95
	rural	2.89	3.09	3.56	2.43	3.87

Table 2.9. Average affective scale ratings for each actual region.
Black outline indicates the highest average rating for each scale, gray shading indicates the second-lowest average rating, and black shading indicates the lowest average rating.

CVS indices did not correlate significantly with any scale ratings, but significant negative correlations were found between *bad* raising indices and *unaccented*, *speaks like me*, *familiar*, and *Californian* ratings ($r_{familiar} = -.68$, $p < .01$; $r_{speaks\ like\ me} = -.72$, $r_{unaccented} = -.76$, $r_{Californian} = -.76$, ps $< .005$).

2.2.3. Interaction of identification and language attitudes

Perhaps surprisingly, the average ratings by perceived region in Table 2.10 show several differences from the average ratings by actual region. In particular, the Bay Area gains top ratings on 7 of 14 scales, including all correctness scales except *good English*. Southern

California retains top ratings for self-similarity scales and Californian-ness, and Far Northern California becomes the least prestigious Californian region. The low regard in which listeners hold outside California is even more pronounced in these ratings.

		Perceived region				
	Scales	**Far North**	**Bay Area**	**Cent. Valley**	**So. Cal**	**Outside**
Correctness	*smart*	3.65	4.35	3.90	3.89	3.45
	educated	3.91	4.56	4.06	4.34	3.51
	confident	3.50	4.23	3.77	3.85	3.31
	good English	4.09	4.62	4.16	4.63	3.80
	fast	3.43	3.76	3.57	3.54	2.47
	unaccented	3.89	4.23	3.75	4.09	2.97
Pleasantness	*attractive*	3.24	3.92	3.47	4.13	2.98
	friendly	3.58	4.39	3.89	4.5	3.48
	polite	3.73	4.52	3.99	4.43	3.74
	likeable	3.51	4.34	3.71	4.2	3.51
Self-similarity	*speaks like me*	3	3.57	2.89	3.63	2.17
	familiar	3.83	4.16	3.70	4.36	3.04
Place	*Californian*	4.18	4.50	3.94	4.93	2.71
	rural	3.36	2.66	3.47	2.65	3.93

Table 2.10. Average affective scale ratings for each perceived region.
Cell shading/outlining as in Table 2.9.

An important result from Williams et al. (1999) was that listeners were more likely to identify likeable speakers as a member of their own regional group via the process of *claiming* (see 1.4.3. Dialect recognition, above). In the present data set, Pearson correlations between ingroup identification rates and affective scale ratings (displayed in Table 2.11) suggest claiming effects for Bay Area and Southern California listeners. Listeners from both regions rated speakers who they perceived to be from their regional ingroup higher on *Californian* and *speaks like me*. In addition, Bay Area listeners claimed the speakers who they rated as more *attractive*, and Southern California listeners claimed the speakers who they rated as more *familiar*.

Correctness was apparently not a factor in ingroup identification for either group (with a virtually zero *smart* correlation for Southern California).

	Ingroup		
Scales	**Bay Area**	**So. Cal**	**Outside**
smart	.36	.05	−.43
educated	.39	.30	−.55
confident	.34	.21	−.46
good English	.52	.61	−.67*
fast	.34	.43	−.72**
unaccented	.50	.49	−.94***
attractive	.68*	.61	−.67*
friendly	.49	.51	−.42
polite	.49	.24	−.33
likeable	.55	.23	−.38
speaks like me	.70*	.69*	−.87***
familiar	.58	.70*	−.83***
Californian	.72**	.78**	−.96***
rural	−.41	−.60	.72**

Table 2.11. Correlations between ingroup/outside identification rates (Table 2.8) and affective scale ratings.
"Ingroup" columns pertain to Bay Area or Southern California listeners only. "Outside" column includes all Californian listeners. Asterisks indicate significance (one-tailed t test with df = 10): * $p < .01$, ** $p < .005$, *** $p < .001$.

Californian listeners' outside identification rates correlated significantly with eight attitudes scales. These correlations mostly suggested either a lack of correctness (*bad English*, *slow*, *accented*) or a sense of difference/distinction from the listener (*doesn't speak like me*, *unfamiliar*, *not Californian*). Unlike ingroup identification, attitudes about (in)correctness, especially those relating to language, *were* implicated in outside identification. The significant positive correlation of outside identification with *rural* lends further weight to the idea that listeners associated positive attributes with urbanness (or negative attributes with rurality).

At first blush, it is surprising that the non-Californians, all of whom were from the New Jersey/Long Island suburbs of New York City, were rated high on rurality given the salience of negative linguistic stereotypes around New York City; to Californians, New York City linguistic stereotypes radiate out to encompass the entire East Coast (Fought 2002). The "outside California" write-in data explains this apparent contradiction, as speakers who were specifically identified as being from the East Coast were rated significantly lower on *rural* than those who were identified as being from outside California but not the East Coast (2.94 vs. 4.20, $t_{15} = 2.84$, $p < .01$). (East Coast identification did not significantly affect any other scale.) In other words, the high rurality ratings for non-Californians are not an indication of the densely populated areas of the Northeast US losing an association with urbanness, but rather a function of a lack of detail in identifying where these speakers are from.

2.3. Discussion

It is clear from the correlations in Table 2.11 that there are stronger attitudes associated with *not* being a Californian than being one, as correctness traits such as speaking 'without an accent' are denied to non-Californian speakers even if such traits are not necessarily attached to Californian speakers. This result makes sense given the linguistic security exhibited in PD map-drawing studies by Californians (Fought 2002), as well as Western speakers more generally (Fridland & Bartlett 2006). Consistent with Fought's (2002) hand-drawn maps, these Californian listeners clearly had a notion of linguistic correctness as something non-Californians lack, but did not ascribe correctness to Californians. The negative correlation between *bad* raising and *unaccented* indicates that listeners responded negatively to this non-Californian (and non-Western) feature, although *bad* raising did not significantly affect listeners' responses to any other correctness trait.

The ingroup correlations reveal that self-perceptions differ slightly between California's two most populous regions. Bay Area listeners claim the speakers who they also judge to be most attractive, Californian, and speaking like them; conversely, Southern California listeners claim the speakers who they also judge to be most familiar, Californian, and speaking like them. Interestingly, this creates a dispute between the regions, as each region believes their *own* speakers to be central (and other speakers peripheral) to a notion of sounding Californian. Southern Californians are apparently more comfortable staking a claim to the CVS itself, however, as the two speakers with the highest CVS indices (BA1 and SC1) had much higher ingroup identification for Southern California listeners than Bay Area listeners.

These data are similarly mixed on which region is considered most linguistically prestigious; although Southern Californians are rated highest for correctness, pleasantness, self-similarity, and *Californian* scales, listeners ascribe greater correctness to speakers they believe to be from the Bay Area than those they believe to be from Southern California. This boost in correctness is in line with Bucholtz et al.'s (2008) finding that Northern California (including the Bay Area) is believed to feature the "best" English in California thanks to its educational institutions. More broadly, this result reinforces the general finding that folk beliefs about certain places can create a disconnect between what listeners *assume* they're hearing and what they *actually* hear (e.g., Niedzielski 1999). Bucholtz et al. also found a belief that Southern California has the "worst" English, which conflicts with the positive attitudes toward Southern California speakers found in the pilot study; it is possible that the factors cited as driving stigmatization (slang and ethnolinguistic diversity) were less relevant in the pilot study, as both word choice and ethnicity were invariant between speakers. Indeed, the degree to which a speaker used CVS features apparently had no effect on their affective scale ratings.

The less populous regions, conversely, received less favorable ratings, a result that is perhaps not surprising in a schema where urbanness is a factor in prestige. As mentioned above, Lower Central Valley speakers were apparently confused with non-Californians, which may have something to do with the common erasure of this region (e.g., Almost no one lives here: Bucholtz et al. 2007:338). Despite the fact that most of the state is covered by farmland, forest, or desert, rural speakers are peripheral to a sense of "sounding Californian."

Finally, the recognition results show that Californians encounter substantial difficulty in identifying the regional origin of Californian speakers, as only two of eight Californian speakers were accurately recognized by a plurality of listeners. Californians may be better able to guess at the regional origin of speakers in naturalistic interactions, of course, as naturalistic interactions may feature strong identifiers such as Southern Californians' use of *the* preceding freeway names (Geyer 2001) or Northern Californians' use of *hella* (Fought 2002). While there was variation in Californian and Western vocalic features between the speakers in this sample, it is unclear how much of this variation can be attributed to regional origin versus the host of other social factors that can drive variation.

2.4. Methodological changes for main study

While the pilot study provides a useful early picture of the social meanings of California English, it also suggested several methodological changes that would be needed to adequately address this dissertation's research questions (1.7.1. Research questions). These changes, which were addressed in the main study, were a larger and more geographically diverse listener sample, a less ethnically homogeneous speaker sample, affective scales that better reflected the folk-linguistic framework of the wider speech community, and more sophisticated methods of creating stimuli.

The listener sample greatly underrepresented Far Northern California and the Lower Central Valley, as only 3 of the 41 listeners called either region home. The local preferences found in research on perceptual dialectology (e.g., Preston 1996b) and mental mapping more generally (Gould & White 1986) should lead us to assume that respondents from these regions would not simply assent to their stigmatization at the hands of outsiders. Exactly how these individuals may resist (or decline to resist) their peripheral status is unclear, however, nor is it predictable how these listeners would fare on a recognition task. While the main study eliminated Far Northern California as a region (3.1. Regions), it succeeded in recruiting Lower Central Valley listeners (in fact, the Lower Central Valley contributed more listeners to the main study than either the Bay Area or Southern California). Moreover, the fact that listeners in the pilot were all summer-session students at UC Davis meant that they had all lived for at least a year in the Greater Sacramento region and among students from regions of California beyond their own. In the main study, social-network sampling was used to elicit responses from listeners in more regionally homogeneous communities.

Similarly, a speaker sample featuring only Caucasian native speakers of English is simply too homogeneous to represent the full ethnolinguistic diversity of California (e.g., Table 2.2). Indeed, Bucholtz et al.'s (2007) hand-drawn maps of the state clearly demonstrate that ethnicity is an important driver of Californians' sense of what Californians sound like, and variationist research by Fought (1999) and Eckert (2008b) shows that California sound changes belong as much to the state's minority speakers as its Caucasian speakers. The main study addressed this need by incorporating ethnicity as a speaker factor, recruiting both Latinas/os and Caucasians for inclusion in the stimulus pool.

The affective scales in the pilot study were selected based on "correctness" and "pleasantness" constructs that past perceptual dialectology research has found to be salient in other parts of the US (e.g., the Upper Midwest). As discussed in the Introduction (1.6.2. California perceptual dialectology), it is questionable whether Californians share even the same underlying framework of "standardness" as Upper Midwesterners (Fought 2002), raising the issue of whether the scales used in the pilot adequately reflected Californians' folk-linguistic framework. Instead of imposing pre-existing language attitudes constructs, the majority of the scales used in the main study were drawn from a pretesting task in which Californian listeners generated affective traits to describe speakers.

The presence of interspeaker variation in the pilot stimulus sample makes it difficult to tease out the effect of vocalic variation on either identification or attitudinal responses from the effect of other types of variation. Speaker BA1, for example, had by far the most California-shifted vowels in the sample sentences, which may have contributed to her 0% ingroup identification and low correctness ratings. On the other hand, SC1 was the second-most shifted but enjoyed both positive ratings and the highest ingroup identification rate. In short, the pilot results offer only indirect hints about the social meaning of California Vowel Shift features. In order to address this need, the main study used matched-guise stimuli differing only by the use of two CVS features: TRAP backing and GOOSE fronting.

Finally, in deciding how to create stimuli for the pilot, I opted for highly controlled conditions (sentences read in a phonetics booth), the idea being that by using content that was inherently evaluatively 'neutral,' listeners' reactions would be responsive only to speakers' accents—that is, that 'neutral content' was necessary to ensure the generalizability of the task's results. For reasons I outline in the Introduction (1.4.1. Language attitudes), however, I argue

that ‘neutral content’ is not only unnecessary for stimulus generalizability, but it may well be impossible to create stimuli that are truly neutral in content (Campbell-Kibler 2006:82–85; Campbell-Kibler 2009). Moreover, a long-established principle in variationist sociolinguistics—that greater attention to speech causes speakers to style-shift toward more standard variants (Labov 1972c)—suggests that this task’s stimuli probably did not capture the full extent of interregional differences given the circumstances of their recording; indeed, previous studies have shown that California speakers do utilize CVS features stylistically (1.6.1. California indexicality). As a result, the main study elicited stimuli via a cartoon retell task, which generated spontaneous speech while limiting topical variation. This retell task was not a standalone task, but rather was integrated into a larger sociolinguistic interview structure, in order to increase speakers’ comfort and willingness to use more casual, naturalistic speech in the retell task.

Chapter 3
Methods

Below are my research questions:

1. To what extent (if any) do Californians perceive the California Vowel Shift as Californian?
2. What evaluations do Californians attach to the California Vowel Shift? How do these evaluations compare to popular portrayals of California speakers?
3. To what extent (if any) do listeners from different regions of California evaluate speakers of California English differently?

In order to investigate these research questions, I conducted a dialect recognition task in which listeners heard samples of spontaneously produced speech, guessed where speakers were from, and rated speakers on language attitudes scales. Each stimulus in this task represented either a California-shifted guise or a conservative (non-shifted) guise. A set of scripts coded in

Praat (Boersma & Weenink 2015) created these acoustically manipulated guises via vowel resynthesis.

For the purpose of creating these guises, the California Vowel Shift (CVS) was operationalized by focusing on two of its features: TRAP backing and GOOSE fronting. These features were selected for several reasons. First, each stands in for a major subsystem of the CVS, with TRAP backing representing the backing and/or lowering of front lax vowels and GOOSE fronting representing the fronting of high/mid back vowels. Importantly, TRAP backing and GOOSE fronting also tend to pattern together.[1] Second, although neither feature is unique to California, especially GOOSE fronting,[2] popular portrayals of California English have incorporated both of these features (Pratt & D'Onofrio 2014), so they are arguably associated with California (and Californian personae) at some level of public awareness. Third, whereas most studies on the CVS have ignored speaker ethnicity and/or investigated primarily Caucasian speakers, Eckert (2008b) and Fought (1999) have shown that Chicana/o speakers also use TRAP backing and GOOSE fronting to index their place in the social landscape (see 1.6.1. California indexicality, above).[3]

Prior to the dialect recognition task, I carried out sociolinguistic interviews and a pretesting task. These stages were designed with two goals in mind: in the short term, facilitating the main dialect recognition task, and in the long term, generating production and folk-linguistic

[1] My analysis of Cory Holland's (2014) California data found that speakers' mean Euclidean distance between TRAP and TRAP-N correlated significantly with their mean Euclidean distance between GOOSE and GOOSE-L ($r = .3575$, $t_{62} = 3.01$, $p < .005$).
[2] The Atlas of North American English identifies GOOSE fronting as characteristic of the West, but GOOSE fronting is also found in the Southeastern US, Philadelphia, and Canada (Labov et al. 2006b:101–3). TRAP backing appears to have spread to much of the West (e.g., Becker et al. in press; Fridland & Kendall forthcoming; Holland 2016).
[3] It should be noted that Eckert's (2008b) study did not examine speakers' degree of TRAP backing itself but the TRAP/TRAP-N nasal split, of which TRAP backing is a part.

data that could be analyzed beyond this project. The following sections describe the selection of Californian regions (3.1), the purpose and design of the interviewing stage (3.2) and pretesting stage (3.3), and finally the methods of the main study itself (3.4).

3.1. Regions

In order to conduct interviews, it was first necessary to define the regions of California from which speakers would be drawn, as this would influence the design of the dialect recognition task. I chose three regions of California: the (San Francisco) Bay Area, the Lower Central Valley, and Southern California (Figure 3.1). Although the existing evidence for intrastate differences in California English is minimal (see 1.3.1. Intrastate variation in California English, above), these regions enjoy different folk-linguistic status (see 1.6.2. California perceptual dialectology, above). Southern California linguistic stereotypes tend to stand in for the state as a whole (e.g., Podesva 2011; Preston 1996b), and Californians themselves construct both human-geographic and folk-linguistic difference along a north vs. south dichotomy (Bucholtz et al. 2008; Fought 2002; Montello et al. 2014), often erasing the less-urbanized Lower Central Valley (Bucholtz et al. 2007). As the Pilot Study chapter details, these regions also differ along lines of sociocultural factors that are often responsible for dialect differentiation (Wolfram & Schilling-Estes 2006:29–35): economic ecology, geography, settlement, and language contact (see 2.1.1. Regions, above).

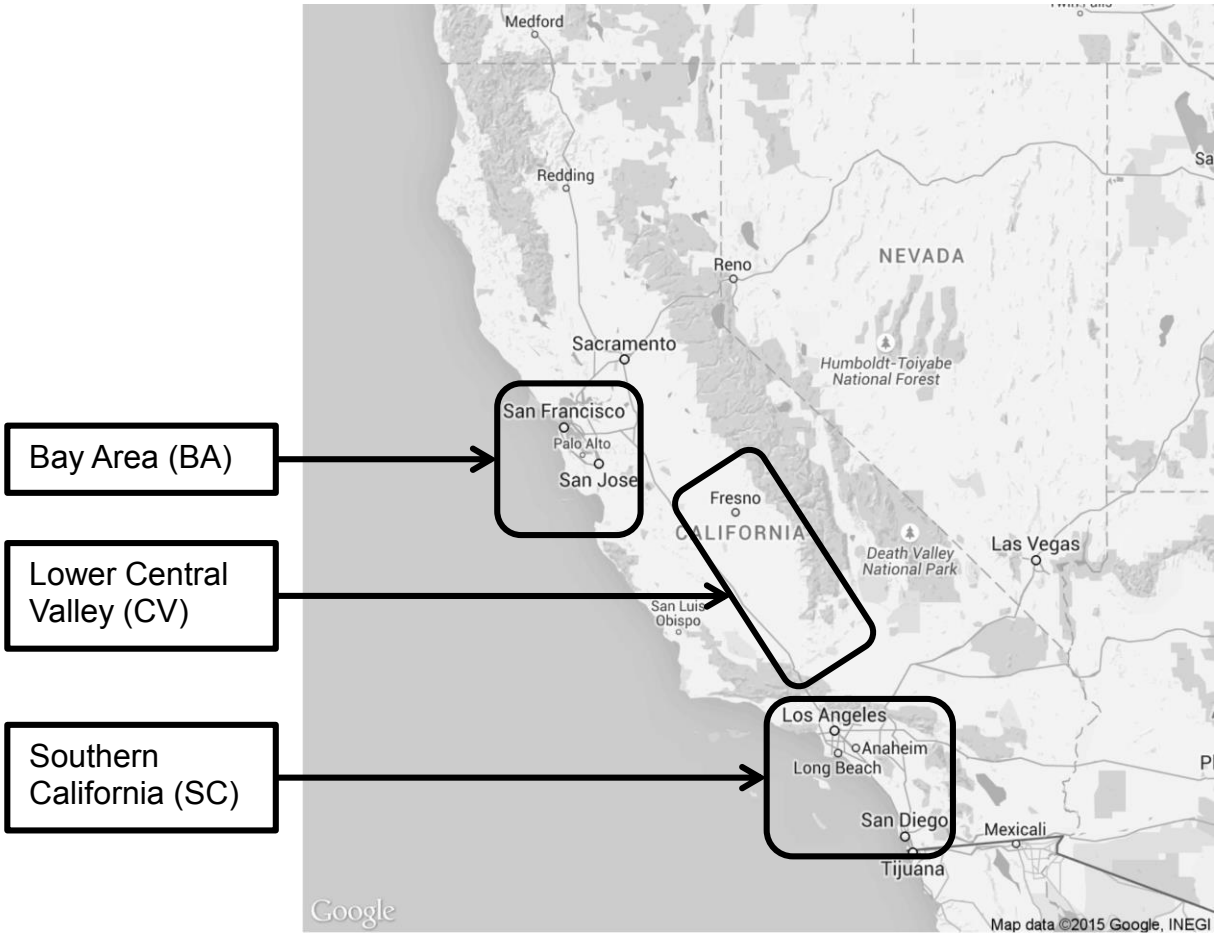


Figure 3.1. Californian regions and abbreviations in the main study.

These regions excluded some parts of the state (e.g., Sierra Nevada mountains, Mojave Desert), most of which are sparsely populated in comparison to the three focal regions; although the focal regions encompass a minority of the state's land area, they account for 84.4% of California's population.[4] As in the pilot study, the heavily populated Greater Sacramento region was excluded because of the ambiguity of its human-geographic and folk-linguistic position; as a result, the Lower Central Valley was defined to exclude areas in the southern vicinity of Sacramento. The only modification to the pilot study's regions was the exclusion of Far Northern

[4] This figure is based on US Census Bureau 2014 county-level population estimates, available at http://www.census.gov/popest/data/counties/totals/2014/index.html. The total population of these three regions is 32,740,467, whereas California's total population is 38,802,500.

California, defined as the entire part of California above the latitude line connecting Lake Tahoe and Clear Lake. Compared to the other regions, Far Northern California does not seem to have much of a consistent folk-geographic or folk-linguistic profile. Moreover, the relative lack of ethnic diversity in Far Northern California would make it difficult to fill the desired demographic cells for this region, as I planned to include both Latina/o and Caucasian speakers in the speaker sample.

3.2. Interviews

I conducted 30 sociolinguistic interviews between March and May 2014 in three research sites selected to represent the three focal Californian regions: Fresno (which represented the Lower Central Valley), Orange County (which represented Southern California), and various locales within the San Francisco Bay Area. As mentioned above, the primary goal of the sociolinguistic interviews was to produce stimuli for the main dialect recognition task; this goal was met through the inclusion of a cartoon retell task within the interview structure. Whereas in the pilot study, stimuli were drawn from isolated recordings of a pre-determined list of sentences (see 2.1.2. Speakers and stimuli, above), the retell task was included to allow for speech that was not only more natural, but also varying minimally in content from one stimulus to the next. The secondary goal of the interviews was to generate production and folk-linguistic data that will serve as the basis for future analyses. As such, the sociolinguistic interviews in this study were designed along lines of classic sociolinguistic research (e.g., Labov 1984), comprising a long conversation section followed by several shorter tasks: a cartoon retell, a reading passage, a word list, and minimal pairs. Interview tasks were always conducted in this order, though for a handful of speakers one or two tasks were omitted for the sake of time.

Since the retell task represented the crucial element of these interviews, I discuss the design of the retell task first (3.2.1) before describing the design of the rest of the interview (3.2.2). I then describe speaker recruitment and interview setup (3.2.3), I briefly profile the speakers (3.2.4), and I describe final steps before the pretesting task (3.2.5).

3.2.1. Retell task design

The retell task responded to the need for stimuli that fulfilled certain criteria. First, stimuli had to be spontaneously produced in a relatively casual speech style so as to approximate speakers' vernaculars. Since more formal speech styles move speakers toward more standard speech and away from their vernaculars (Labov 1972d), spontaneously produced stimuli would encompass a broader range of interspeaker variation than would be captured by a read-speech elicitation task such as a word list. Second, while I argue that stimulus content can never be truly 'neutral' (see 1.4.1. Language attitudes, above), the content of the stimuli needed to avoid overly shading listeners' perceptions. For example, a speaker talking about getting a perfect SAT score would probably receive high *smart* and *educated* scale ratings regardless of accent. Third, the stimuli had to contain a sufficient number of TRAP and GOOSE tokens since these vowels would be manipulated to create the matched guises for the main task.

In the retell task, speakers watched short cartoon videos and then recounted the story in their own words, as if speaking to a friend who had not seen the video before. This task used two videos from a series of online cartoons called "Simon's Cat," created by British animator Simon Tofield (https://simonscat.com/); these videos feature the hijinks of a male human character, Simon, and his anthropomorphized, mischievous unnamed cat. In order to seed TRAP and GOOSE

tokens, I renamed the cat "Matt the cat" and his owner "Stu."[5] (In the retell task, interviewees were given a card with the characters' names and pictures so they could remember them.) The first video, "Scary Legs," depicts Matt the cat's attempts to subdue and warn his arachnophobic owner about a spider (Figure 3.2). The second video, "Suitcase," depicts Matt the cat interfering with Stu's attempts to pack for a trip (Figure 3.3).

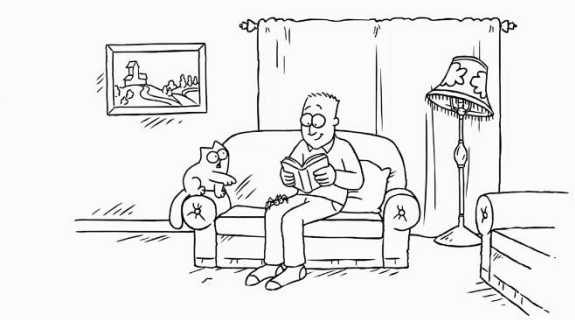


Figure 3.2. Still image from the "Scary Legs" video (https://youtu.be/19jaOSNibkU) used in retell task. Image reprinted with permission.

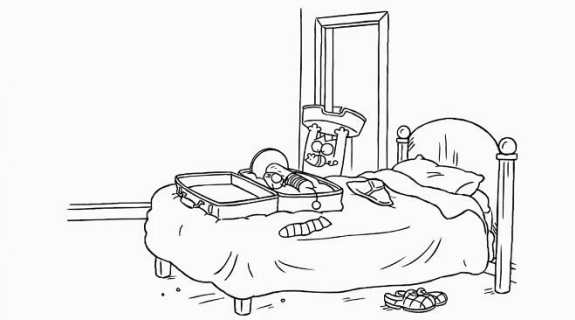


Figure 3.3. Still image from the "Suitcase" video (https://youtu.be/_dm_2G-rIOs) used in retell task. Image reprinted with permission.

At 96 seconds and 105 seconds, these videos were short enough that speakers could remember all of the actions in the videos. At the same time, these videos contained enough individual actions by Matt the cat and Stu to provide sufficiently long retells to draw stimuli from; importantly, the action switches back and forth between Matt the cat and Stu, so speakers have a reason to keep saying the characters' names rather than using pronouns. (The fact that both characters are male also lessens the likelihood of pronoun substitution, as *he* could refer to either character, although in actual discourse it is not uncommon for *he*/*him*/*his* to refer to different antecedents in the same sentence.) The importance of reporting the characters' actions was reinforced by the verbal prompt I gave speakers prior to this task: "Okay, so for this part,

[5] I deliberately sought cartoons starring cats as a way to introduce TRAP tokens, although I avoided well-known cartoon cats such as Sylvester or Felix since speakers would be likely to use the cats' given names rather than calling them "the cat."

I'm going to show you some short, funny cartoons about a guy named Stu and his cat, Matt the cat. After each one, I want you to re-tell the story in the video to me as if you were telling a friend who hadn't seen the video before. Make sure to mention as many of Stu's and Matt the cat's actions in the video as possible for maximum comedic effect."[6] As the prompt suggests, these videos were light-hearted and humorous, which helped to engage speakers in the task and made plausible the context of telling a friend about the videos.

Because speakers tell the story in their own words, this task meets the first goal of eliciting spontaneously produced speech rather than scripted speech in order to better capture interspeaker variation. The content of speakers' retells is mostly limited to the content of the cartoons, which fulfills the second goal of controlling for topical variation in the stimuli; the relative neutrality of excerpts of the retells was later checked by pretesting these excerpts (see 3.3. Pretesting, below). Importantly, limiting topics to the content of the videos also prevented stimuli from containing information that directly gives away speakers' personal characteristics. In order to ensure that these videos would satisfy the third goal of producing a sufficient number of TRAP and GOOSE tokens, I informally pretested this task with friends and students. While there was interspeaker variation in the length and detail of speakers' retells, this pretest demonstrated that the use of character names "Matt the cat" and "Stu," along with the second video revolving around a suitcase, helped to produce a sufficient number of tokens.

In the sociolinguistic interviews themselves, the retell task was successful in eliciting a sufficient quantity of speech from which to draw excerpts for stimuli. Retells ranged in duration

[6] This prompt varied somewhat from interview to interview. The request to mention as many actions as possible was almost always present.

from 38 seconds to 214 seconds, averaging 89 seconds;[7] the average "Scary Legs" retell, at 93 seconds, was longer than the average "Suitcase" retell, at 78 seconds. Most retells produced a sufficient number of TRAP and GOOSE tokens, ranging from 6–48 TRAP tokens (median: 18 tokens) and 1–25 GOOSE tokens (median: 8.5 tokens). As I had anticipated, *cat* and *Matt* were the most frequent TRAP tokens, and *Stu*(*'s*) and *suitcase* were the most frequent GOOSE tokens.[8]

3.2.2. Conversation and task design

The retell task represented only a short portion of the longer sociolinguistic interviews, which also included a long conversation section, a reading passage task, a word list task, and a minimal pairs task. The conversation section was designed with two goals in mind: first, increasing speakers' familiarity and comfort with the interview context, in order to encourage more natural speech in the retell task, and second, gathering a large body of production and folk-linguistic data for future analyses. The reading passage, word list, and minimal pairs tasks were designed to elicit speakers' production in more formal speech styles; a future analysis of this data in combination with production data from the conversation section will focus on whether CVS features are sensitive to attention paid to speech. As this dissertation focuses only on the data produced by the retell task and none of the other sections of the interview, however, I will describe these sections of the interview only in brief.

The conversation section was designed along lines of classic sociolinguistic interviews, which gather relatively natural speech by lessening the attention that speakers pay to their own

[7] These statistics are based on the 17 retells that were transcribed in summer 2014 (see 3.2.5. Post-interview process, below).
[8] The majority of GOOSE tokens (57 of 61) were postcoronal, including the most frequent words *Stu*(*'s*) and *suitcase*, which is important given that the postcoronal environment promotes GOOSE fronting (Labov 2010; Labov et al. 2006b). The calculation of manipulation targets took phonetic environment into account by adjusting the targets for non-postcoronal tokens (see 3.4.2. Manipulation target calculation, below).

speech (Labov 1984; Schilling 2013). In this study, I framed interviews by positioning myself as an outsider to California who was curious about the experience of growing up and living in different parts of the state. (I grew up in Delaware and spent little time in California prior to my graduate career at UC Davis.) I chose this positionality for several reasons. I sought to reverse the asymmetric power relationship inherent in sociolinguistic interviews by positioning my interviewees as experts. At the same time, this positionality emphasized that these interviews were not for the purpose of gathering hard facts per se (and thus interviewees would not be held accountable for the factual value of what they said), but a matter of satisfying personal curiosity. This framing guided the interview questions that I wrote, such as questions about childhood and adolescence that especially aimed at eliciting personal narratives. Moreover, given that I am not a member of the California English-speaking speech community, I felt it was necessary not to fake membership in it.[9]

I attempted to let speakers guide the conversation by engaging them on the topics that appeared most interesting to them and eliciting narratives where possible. This meant asking follow-up questions and maintaining topical coherence rather than changing topic by asking unrelated, abrupt questions. Prior to conducting interviews, I wrote a list of general questions (Appendix A) that I could use to stimulate conversation, but I was happy to abandon my pre-written questions in interviews in favor of more interesting topics that speakers brought up. The exception was a series of questions that were designed to elicit folk-linguistic discourses around California English, including popular portrayals and stereotypes thereof, which I brought up

[9] One potential drawback to this positionality is that it may have caused speakers to accommodate to my non-Californian accent (e.g., Coupland 1984). Then again, it is possible that some speakers would do so even without me explicitly marking myself as an outsider.

myself at the end of the conversation section if folk-linguistic topics did not arise naturally. These questions drew upon the interview methods of Niedzielski and Preston (2003).

For the reading passage task, speakers read a modified version of "Comma gets a cure"[10,11] (Honorof et al. 2000). Speakers were asked to read the passage to themselves once before reading it aloud at their own pace. Speakers were not given instructions about how to treat the reading passage as a discourse event (e.g., "read it naturally"). The word list consisted of 105 monosyllabic or disyllabic words representing 11 vowels in numerous phonetic environments. The word list was presented in the same randomized order to all speakers, with words rearranged such that minimal pairs did not appear as adjacent words in the list. In the minimal pairs task, speakers were asked to read aloud pairs of words and state whether they believed their pronunciation of the words was the same or different. The presence of several obvious homophones like *bred*/*bread* and obvious non-homophones like *dude*/*did* meant that speakers could not assume that all pairs were meant to sound the same or different; speakers were informed that a number of pairs would be obvious. The minimal pairs list was presented in the same randomized order to all speakers. The text of the reading passage, word list, and minimal pairs task can be found in Appendix B.

After the minimal pairs task, the recording was stopped and speakers were given a paper demographic questionnaire (Appendix C). This questionnaire generally took speakers five

[10] Comma Gets a Cure and derivative works may be used freely for any purpose without special permission, provided the present sentence and the following copyright notification accompany the passage in print, if reproduced in print, and in audio format in the case of a sound recording: Copyright 2000 Douglas N. Honorof, Jill McCullough & Barbara Somerville. All rights reserved.

[11] This reading passage was modified to introduce a greater number of Californian/Western features, especially prelateral tense/lax mergers (e.g., FLEECE-L/KIT-L); these modifications were partially based on Holland's (2014) modifications. Several phrases that struck me as unnatural were also revised (e.g., "a deserted district of the territory").

minutes or less to fill out. I initially planned to administer the questionnaire verbally but realized that using a paper questionnaire would give me time to save the sound files and disassemble the recording setup to make a cleaner exit from the interview location. Before I left, I thanked the speaker for participating, asked if they wanted any information deleted (few did) in keeping with common ethical practice (Schilling 2013), and invited any additional questions.

No matter interviewers' efforts, the relatively structured and formal nature of one-off sociolinguistic interviews means that interviews seldom truly obtain speakers' vernaculars (Schilling 2013), and this research was no different.[12] This built-in limitation notwithstanding, my general impression was that most of the interviews were highly successful in obtaining relatively casual speech—if not necessarily speakers' vernaculars—with some speakers making unsolicited comments that the conversation felt casual and enjoyable to them. Interviews averaged 96 minutes in duration, with the conversations averaging 77 minutes. Most of the variability in interview duration was due to variability in the duration of the conversation; the shortest conversation ran just 28 minutes (six lasted less than an hour), and two particularly lively conversations breached the two-hour mark. The reading passage, word list, and minimal pairs tasks all went as planned, averaging 11 minutes in total duration. In total, the 30 interviews yielded 48 hours, 11 minutes of data.

3.2.3. Speaker recruitment and interview setup

I conducted thirty interviews across a three-month period in 2014 in three interview sites: Fresno (representing the Lower Central Valley), Orange County (representing Southern California), and various locales within the San Francisco Bay Area. Six speakers were

[12] My interviews contain virtually no casual speech in what Labov (2006:67) labels "Context A_2," speaking to someone else in the room (possibly because in setting up the interview, I specifically recommended a private location).

interviewed in Fresno in March, four speakers were interviewed in San Jose in April, ten speakers were interviewed in Orange County in May, and ten speakers were interviewed in San Francisco and the East Bay (i.e., Alameda and Contra Costa Counties) in May. Due to constraints around my teaching schedule, these interviewing trips were relatively short (four to six days).

I selected these sites primarily based on the locations of California State University (CSU) campuses whose linguistics professors aided in speaker recruitment.[13] A professor at Fresno State University advertised the study to his introductory linguistics course and his Masters students; a professor at CSU Fullerton (in Orange County) allowed me to give a research presentation to a graduate course and advertised the study to her undergraduate historical linguistics course; and an anthropology professor at CSU East Bay advertised the study to anthropology majors, graduate students, and students in a language and culture course. In addition, after a trip to San Jose that was relatively unsuccessful in terms of recruiting, I increased the level to which I solicited the help of Californian friends and former students in making initial contacts with speakers in Southern California and (especially) the Bay Area. Beyond these initial points of contact, I also sought to recruit speakers through "friend-of-friend" sampling, whereby speakers put me in contact with friends or associates who could become additional interviewees. I reasoned that, given the relatively heavy time commitment of participation (90 minutes) and lack of material incentive for participation,[14] prospective speakers would be more likely to volunteer if a friend who had already participated vouched for me. In all,

[13] CSUs were chosen for recruitment because, as compared to University of California (UC) campuses, they are more likely to reflect the ethnic diversity of the communities in which they are geographically situated, and students are more likely to be local. In addition, although I recruited speakers primarily through linguistics professors, only some of my speakers were linguistics students.

[14] No interviewees received monetary or material compensation, although some professors offered extra credit to their students for participation.

14 speakers were recruited through contacts with professors, six through contacts with friends, six through contacts with students, and just four through contacts with other speakers. Aside from briefly meeting two interviewees a few months before their interviews, all interviewees were strangers to me.

Prior to interviewing, I set criteria for speakers to be included in the stimulus pool (i.e., speakers from which stimuli for the pretesting and main task would be drawn): being between 18 and 30 years old (inclusive), having lived only in the same region of California for their entire lives, identifying as Latina/o or Caucasian, speaking English natively, and having no speech or hearing impediment. At first, I only recruited speakers who strictly fit these criteria in the name of saving interviewing time. After having to turn down several prospective speakers in my first interviewing trip (in Fresno), however, I realized that interviewing more speakers would aid "friend-of-friend" recruitment—a speaker who did not fit all the stimulus pool criteria could potentially lead me to a speaker who did—and increase the size of the production data set. As a result, I relaxed my criteria for speaker recruitment—eliminating the ethnicity criterion and upper age limit (the lower limit of 18 was necessary for IRB purposes), and broadening the regional criterion to include speakers who had spent some time outside the region—while retaining the stricter criteria for inclusion in the stimulus pool.

When prospective interviewees emailed me to set up an interview, I invited them to choose an interview location, noting that "it should be somewhere relatively private and quiet (to get the best quality recording possible), but also somewhere you're comfortable (since most of the interview will be a casual conversation, and it's hard to be casual when you're somewhere unfamiliar)." Most Fresno and Orange County interviews were conducted in a library, as the Fresno State and CSU Fullerton libraries had group study rooms. All Bay Area interviews (and

some Fresno and Orange County interviews) were conducted in speakers' house, apartment, or porch, my lodging in town, a café, or a public, uncrowded area on campus.

All interviews were conducted with two microphones: a Shure WH20XLR head-mounted cardioid microphone worn by the speaker and a Blue Snowball microphone set to record omnidirectionally and placed on a floor or table between the speaker and interviewer. The two microphones had complementary functions; the head-mounted cardioid microphone produced high-quality recordings that primarily captured the speaker and only faintly recorded any other speech or noise, and the omnidirectional microphone captured both the speaker and interviewer, as well as any environmental noise.[15] In addition, the use of two microphones meant that I had a backup in case one failed. As Schilling (2013) notes, some sociolinguists argue that lavalier (shirt-mounted) microphones are preferable for sociolinguistic interviews, since the intrusiveness of a head-mounted microphone calls the speaker's attention to the recording; I opted for the head-mounted microphone instead, however, because the nature of the later phonetic analysis would require the highest possible quality recording in order to capture spectral details of speakers' vowels. The head-mounted microphone also afforded extra flexibility as to interview location; although public spaces are not ideal for interviewing due to the risk of environmental noise, the head-mounted microphone did not record most environmental noise.[16]

The head-mounted microphone was connected via XLR cable to a Marantz PMD661 recorder with a −12 dB microphone attenuation, and the omnidirectional microphone was connected via USB cable to my laptop, a Lenovo ThinkPad T410s, and recorded in Audacity

[15] Unfortunately, the Blue Snowball microphone yielded recordings that were unsatisfactorily noisy. I have heard that some linguistics departments stock these microphones in their labs, but I do not recommend them.
[16] In one interview, the sound of a pot being filled with water a few yards away is barely audible in the head-mounted microphone's recording.

(Audacity Team 2012). (The laptop screen was placed out of sight of the speaker until it was used for the retell, reading passage, word list, and minimal pairs tasks.) Both recordings were encoded as single-channel 16-bit PCM (lossless WAV) files at a 44.1 kHz sampling frequency. In the first few minutes of each interview, I manually adjusted the input volume on the Marantz recorder to obtain recording levels that were high without over-recording; however, this adjustment was not always successful, as some interviewees spoke louder or softer over the course of the interview. Figure 3.4 displays one recording setup. After I turned on each microphone, the interview began with a clap so the two recordings could be synchronized later.

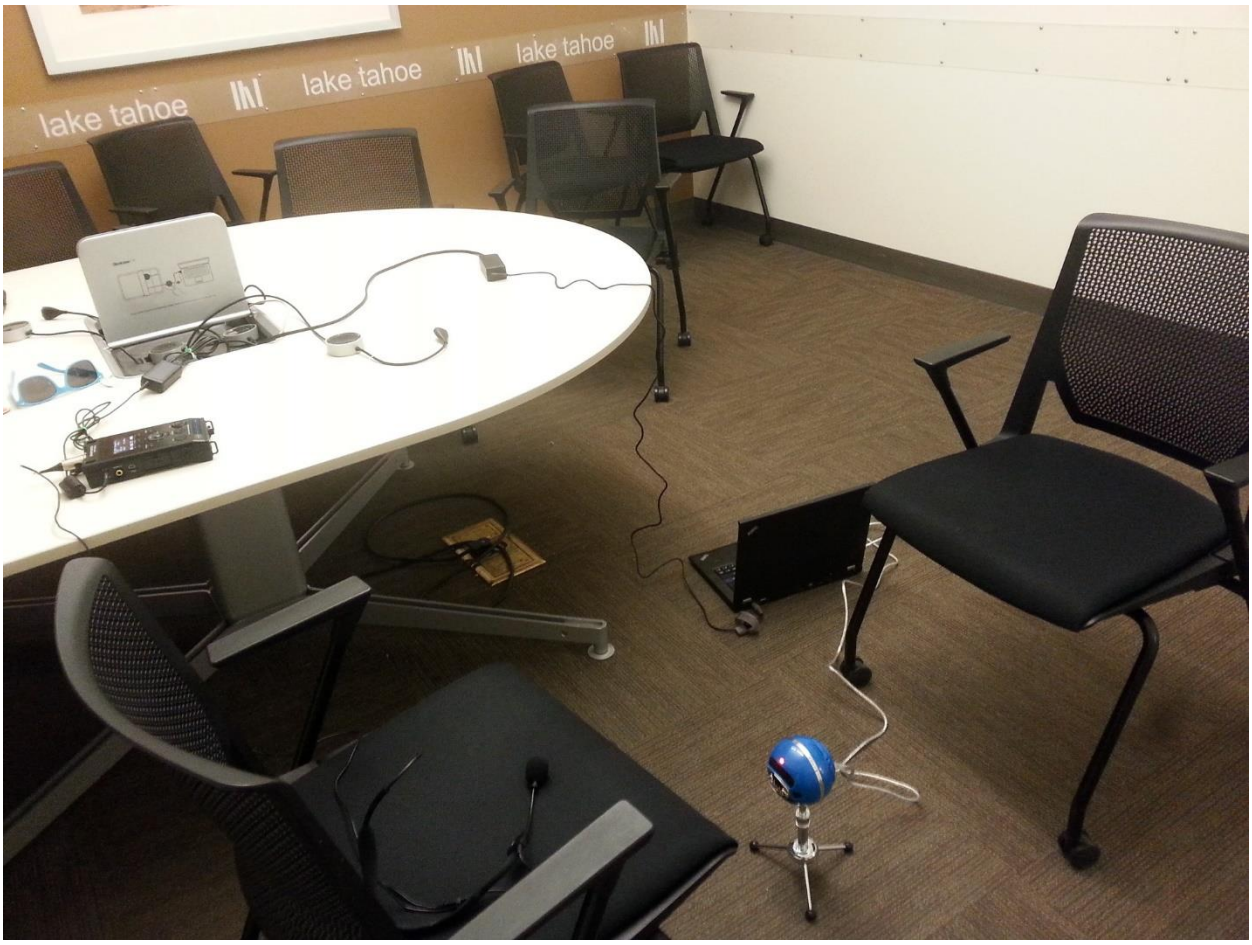


Figure 3.4. A typical recording setup in the Henry Madden Library at Fresno State University.

3.2.4. Speakers

Table 3.1 below displays information on the 30 speakers interviewed in this stage of data collection.

Speaker	Region	Gender	Ethnicity	Age	Hometown
BA01-CH	Bay Area	Male	Caucasian	20	San Jose
BA02-EX	Bay Area	Male	Latino	59	San Jose
BA03-ID	Bay Area	Female	Latina	65	San Jose
BA04-NN	Bay Area	Female	Latina	69	San Jose
BA05-TN	Bay Area	Male	Latino	30	San Pablo
BA06-BB	Bay Area	Male	Filipino	21	Sunnyvale
BA07-SQ	Bay Area	Female	Latina	29	San Francisco
BA08-LD	Bay Area	Female	Latina	25	San Francisco
BA09-FQ	Bay Area	Female	Caucasian	19	Lafayette
BA10-WS	Bay Area	Male	Indian	30	Scotts Valley
BA11-DM	Bay Area	Female	Caucasian	29	Santa Cruz
BA12-UI	Bay Area	Male	Caucasian	20	Pleasanton
BA13-MQ	Bay Area	Female	Caucasian	23	Oakland
BA14-DX	Bay Area	Male	Indian	20	San Jose
CV01-NK	Lower Central Valley	Female	Caucasian	19	Squaw Valley[a]
CV02-DQ	Lower Central Valley	Male	Caucasian	24	Fresno
CV03-NG	Lower Central Valley	Female	Latina	30	Fresno
CV04-CX	Lower Central Valley	Male	Caucasian	19	Lemoore
CV05-FC	Lower Central Valley	Male	Latino	19	Madera
CV06-BQ	Lower Central Valley	Male	Caucasian	22	Fresno
SC01-UO	Southern California	Male	Asian	20	Anaheim
SC02-KH	Southern California	Male	Latino	23	Rowland Heights
SC03-CI	Southern California	Male	Caucasian	81	Los Angeles
SC04-BT	Southern California	Male	Caucasian/Latino	30	Rcho. Cucamonga
SC05-KB	Southern California	Male	Caucasian	26	Orange
SC06-KY	Southern California	Female	Latina	24	Cudahy
SC07-UK[b]	Southern California	Male	Caucasian	24	Chino
SC08-LB	Southern California	Female	Caucasian/Latina	22	Ontario
SC09-DY	Southern California	Female	African American	28	West Covina
SC10-NM	Southern California	Female	Caucasian	20	Orange

Table 3.1. Demographic information on speakers recorded in this study.
Speakers selected for main study stimulus sample highlighted in gray.
[a] CV01-NK's hometown is the Squaw Valley in Fresno County, not the (better-known) Squaw Valley near Lake Tahoe in Placer County.
[b] SC07-UK was selected for the pretesting stimulus sample, but excluded from the main study stimulus sample.

It should be noted that a disproportionate number of speakers had studied linguistics (including some who were near earning Master's degrees in linguistics), especially in Southern California. In the abstract, this is not an ideal situation. For one, linguists make up a small percentage of the population, so my sample overrepresented them. For another, linguists have more linguistic awareness overall (by definition) than folk linguists, and when it comes to the dimension of control, linguists may be more inclined to modify their speech to come off in a favorable light. Some linguists may be aware of sociolinguistic interview techniques—one speaker asked if I planned to ask him a "danger of death" question[17]—which renders some interviewing strategies moot. Finally, some of the discourses I sought to investigate through my folk-linguistic questions are in conflict with the opinions of linguistics as a field, and linguistics students learn early on to rethink these discourses (Bowie 2013). I was surprised (and pleased) to find, however, that the linguistically aware speakers did express folk-linguistic opinions, even in the presence of an interviewer who, as a sociolinguist, they might expect would participate in norm enforcement discouraging these sorts of discourses. Regardless of the effect of having so many linguists and linguistics students in my sample, it was a predictable consequence of my sampling methods; my short interviewing trips precluded the techniques of embedding myself within local communities or developing extensive "friend of friend" interviewee networks, and for a number of speakers, their main incentive for participation was extra credit in a linguistics course.

Table 3.2 below gives average TRAP and GOOSE productions in the retell task by female and male speakers in the main study stimulus sample (highlighted in gray in Table 3.1), as

[17] I did not.

compared to a standard benchmark and a Californian benchmark.[18] Table 3.2 demonstrates that both TRAP backing and GOOSE fronting are advanced among speakers in the main study stimulus sample, as females' production of TRAP, females' production of GOOSE, and males' production of GOOSE in this study are all shifted further away from the standard benchmark than the Californian benchmark; males' TRAP in this study is comparable to the Californian benchmark.

Feature	Gender	Retell task	Standard	Californian
TRAP	Female	1756	2048	1810
	Male	1614	1731	1601
GOOSE	Female	1945	1374	1700
	Male	1665	1190	1417

Table 3.2. Average TRAP and GOOSE F2 frequencies in the retell task among main study stimulus speakers, compared with benchmarks from previous studies.
All values in Hz. Retell task measurements taken at the vowel midpoint: 438 TRAP tokens (171 from female speakers, 267 from male speakers) and 215 GOOSE tokens (86 from female speakers, 129 from male speakers). Standard values are from Kent and Read's (2002:111–12) averages. Californian values are from Hagiwara (1997:656).

3.2.5. Post-interview process

Once interviews were completed, I listened through interview sound files to delete any sections that could place the speaker at risk of legal harm (e.g., descriptions of past illegal behavior) or direct identification (e.g., the speaker's name, the speaker's spouse's name, the speaker's street address), as well as any information the speaker specifically asked me to delete. Once the files were cleaned, they were divided into sections for the retell, reading passage, word list, and minimal pairs tasks.

The retells for 17 speakers (including all of the speakers who were eligible for the stimulus pool) were transcribed using Transcriber (Barras 2002) by an undergraduate linguistics

[18] Here and throughout, F2 measurements are taken at the vowel midpoint.

major working under my supervision. Since the retells were ultimately to be processed by the Forced Alignment & Vowel Extraction (FAVE) forced phonetic aligner (Rosenfelder et al. 2011), they were transcribed according to the preferred FAVE transcription method. Once I checked and corrected transcripts, they were converted to Praat TextGrids using a script available on the Oregon LVC Lab website (http://lvc.uoregon.edu/trans_to_praat.php). FAVE was then used to produce phonetically aligned TextGrids for each retell.

3.3. Pretesting

The pretesting stage comprised an online survey conducted in October 2014 in which listeners at UC Davis generated open-ended descriptions of interviewees based on short excerpts from interviewees' retell tasks; main study stimuli would later be created from most of these excerpts. The pretesting stage was designed with two goals in mind: first, generating a list of affective traits that listeners would use in the main task to rate speakers, and second, ensuring that the excerpts from which I would create main study stimuli were relatively affectively 'neutral' (i.e., pretesting the excerpts themselves), with the usual caveat that true neutrality is unattainable (see 1.4.1. Language attitudes, above). The pretesting stimulus pool included excerpts from 13 of the 30 interviewees, which was later narrowed to 12 interviewees for the main study stimulus pool (see Table 3.1, above). Two excerpts were chosen for each speaker and edited to remove disfluencies where possible.

The following sections describe the processes of selecting speakers for the pretesting stage (3.3.1), selecting and preparing excerpts from these speakers' retells (3.3.2), and designing the pretesting task (3.3.3); finally, I discuss the results of this task and its implications for the main study (3.3.4).

3.3.1. Speaker selection

In order to be included in the stimulus sample, speakers had to be between 18 and 30 years old (inclusive), have lived in their region their entire life, and identify as either Latina/o or Caucasian. These criteria left 15 speakers, with nine demographic cells filled by one speaker and three cells filled by two speakers (Bay Area female Latinas, Lower Central Valley male Caucasians, and Southern California male Caucasians). Two of the three 'extra' speakers were excluded from the stimulus sample for linguistic reasons. One Bay Area female Latina speaker (BA08-LD) had more CVS features than the other (BA07-SQ), and was thus excluded for being less characteristic of Latina/o speech.[19] One Lower Central Valley male Caucasian speaker (CV06-BQ) had fewer CVS features than the other (CV04-CX), and was thus excluded for being less characteristic of Caucasian speech. Neither of the Southern California male Caucasian speakers (SC05-KB and SC07-UK) could be excluded for linguistic reasons, so both were kept in the stimulus sample for the pretesting task; SC07-UK was later excluded from the main study stimulus sample for reasons relating to listener recruitment.

3.3.2. Excerpt selection and preparation

Once the list of speakers was narrowed down, I selected two excerpts for each speaker that met the following criteria: each had to contain at least two tokens of TRAP and at least two tokens of GOOSE, each had to be around 10–12 seconds long, each had to pertain to the cartoon retell task itself (e.g., an excerpt could not consist of the speaker talking about their own cat), each had to be plausible as a syntactic and prosodic stand-alone unit (e.g., an excerpt could not

[19] To be clear, I am not suggesting that Latina/o speakers in California are necessarily less likely to use CVS features, that Caucasian speakers are necessarily more likely to use CVS features, or that the CVS may index only Caucasian identity (see 1.6.1. California indexicality, above). However, Eckert (2008b) argues that the TRAP nasal split indexes Caucasian ethnic identity, in addition to other identities or stances, in a way that it does not index Latina/o identity. In addition, Holland (2014) found that English speakers' TRAP nasal split was greater than Spanish speakers' (Holland did not investigate ethnicity directly, as it was a confounding factor for home language).

begin or end in the middle of a verb phrase), and each had to be relatively evaluatively neutral so they would not obscure the effect of the matched-guise vowel manipulation in the main study (see the discussion of generalizability in 1.4.1. Language attitudes, above).

I coded and ran a Praat script that read the time-aligned TextGrids for each retell and copied intervals for words containing TRAP and GOOSE onto new tiers, thus allowing for easy visual identification of the location of concentrated tokens (Figure 3.5). Words identified as containing TRAP or GOOSE were checked by hand to ensure that each was a valid, nonreduced token (e.g., *to* could be [tʰuː], a GOOSE token, or [tʰə], not a GOOSE token) and to correct any misalignments.

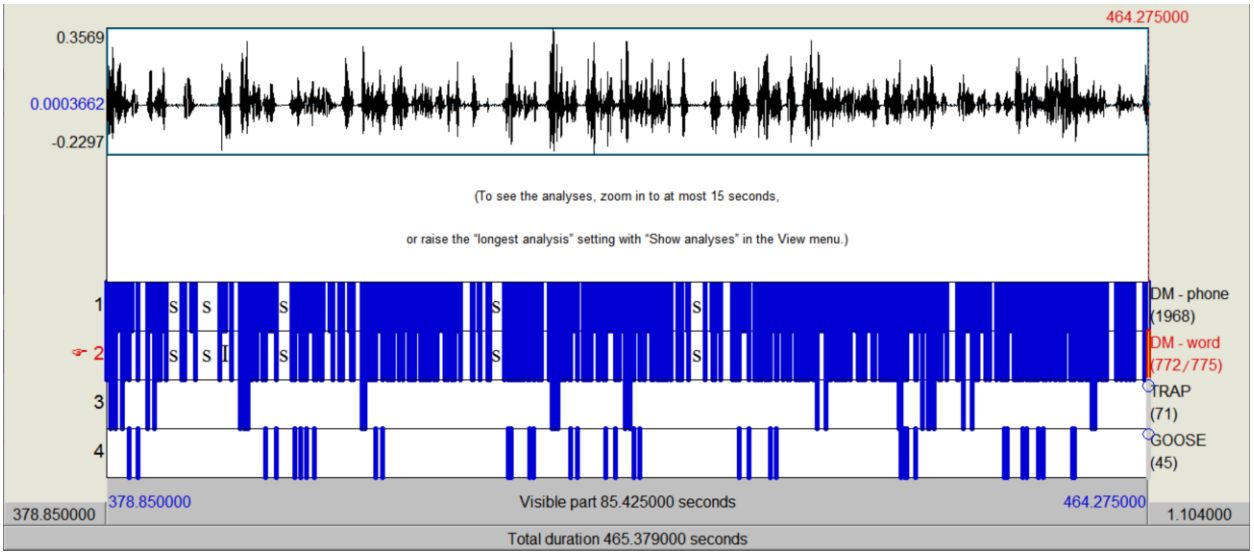


Figure 3.5. Screen capture of BA11-DM's "Suitcase" retell in Praat, demonstrating the location of 15 TRAP and 18 GOOSE tokens in the retell.

Using these time-aligned TextGrids, I identified 80 possible excerpts among the 13 stimulus speakers' retells. It was far more difficult to find excerpts for some speakers than for others, depending on how long the speaker's retell was, how likely they were to use characters'

names rather than *he*, and especially their fluency in performing the retell task;[20] shorter retells featured fewer tokens, and less-fluent retells featured fewer stand-alone prosodic phrases. The set of 80 possible excerpts was narrowed down to 26 excerpts (two per speaker) based on the criteria mentioned above: number of TRAP and GOOSE tokens (with preference given to greater numbers of tokens overall and tokens balanced between TRAP and GOOSE), duration (with preference for excerpts in the 10–12 second range), relevance to the cartoon topic, ability to stand alone as a phrase, and relative evaluative neutrality.

The issue of fluency jeopardized the evaluative neutrality of the excerpts by introducing fluency as a potential extraneous factor in listeners' ratings of speakers. In order to improve fluency, 16 of 26 excerpts were edited to remove long pauses, filled pauses, restarts, partial words, and non-speech sounds (e.g., cough, lip smack). I took care to ensure that the editing process did not result in sudden "jumps" in pitch or loudness or abrupt shifts in rhythm that would make the excerpts sound artificially edited (Figure 3.6). Compare the following excerpt, Exc15, before and after editing, with the edited-out portions highlighted in gray:[21]

> Stu's just sittin' on the couch readin' a book, [0.7] [lip smack] and Matt the cat's just sittin' on the couch like on the, on the edge, [sniff] and uh Stu sees sees the spider crawlin' around he starts freakin' out and he starts pointin' at it, and [0.6] [lip smack] um he try, he's tryin' to tell like Matt the cat to like get it.

> Stu's just sittin' on the couch readin' a book, [0.7] and Matt the cat's just sittin' on the couch like on the edge and uh Stu sees the spider crawlin' around he starts freakin' out and he starts pointin' at it, and [0.6] he's tryin' to tell like Matt the cat to like get it.

[20] Fluency in the retells (which was often unrelated to how fluently the speaker spoke in the conversation section of the interview) was a site of variation among speakers. Whereas some speakers were able to recount the videos' action sequences with ease, others were more likely to use pauses or filled pauses (e.g., *um*, *uh*) in their retells (perhaps due to working memory, interest in the task, or a number of other factors).

[21] Exc15 was edited down from 14.58 seconds to 11.73 seconds; with 2.85 seconds removed, this was one of the more heavily edited excerpts.

In this excerpt, as in others, it was not possible to eliminate all disfluencies without the excerpts sounding artificially edited; in the second line, for example, the "uh" in "and uh Stu" could not be edited out without also removing "and" (which would sound syntactically odd) or disrupting the rhythm of the sentence. In order to check the naturalness of edited excerpts, three linguistically trained listeners (including one phonetically trained listener) listened to excerpts and were asked to identify whether or not they were edited; these listeners could not consistently distinguish the edited excerpts from the non-edited excerpts.

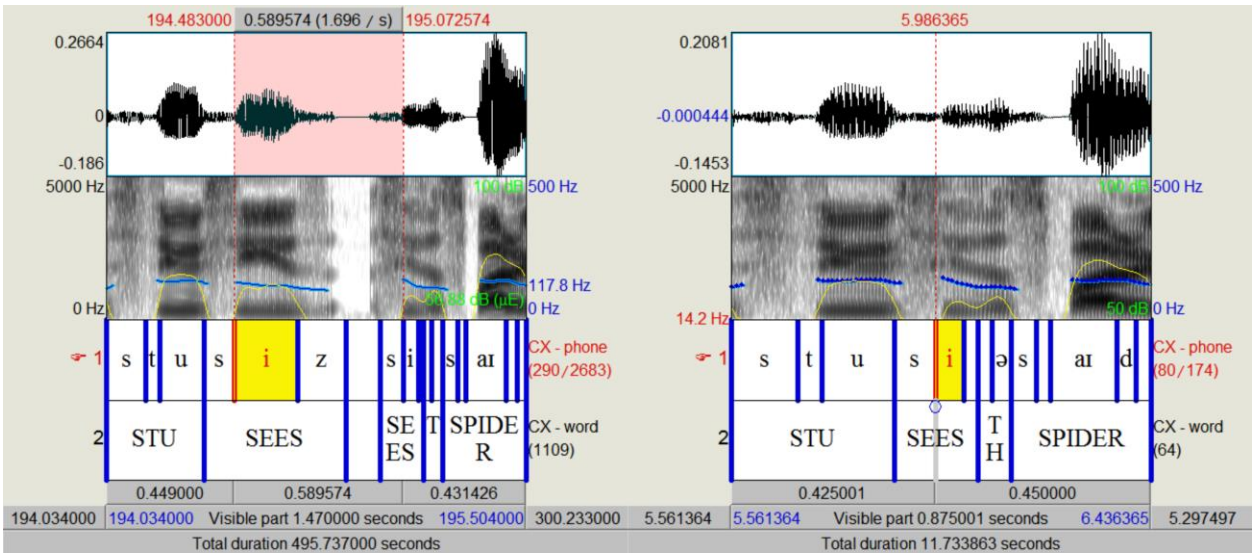


Figure 3.6. Screen captures of Praat sound editor windows before and after editing Exc15 to eliminate the repeated word "sees."
The blue line shows pitch and the yellow line shows intensity. The selected interval in the left panel was eliminated at the point of the cursor in the right panel. Note that in this case, cutting [iz] from the first "sees" and [s] from the second "sees" sounded more natural than simply cutting one of the repetitions.

After excerpts were selected and edited, each excerpt was set to an average intensity of 60 dB. Sound files were converted to .mp3 format in order to be uploaded to the online pretesting survey.

3.3.3. Pretesting survey design

The pretesting task was designed as an online survey hosted by Qualtrics (https://www.qualtrics.com/) licensed to the UC Davis Division of Social Sciences. Given that its primary purpose was to generate affective traits, the task was open-ended in nature. All participants gave informed consent prior to participation.

In each trial, listeners were presented with a single stimulus (which they could click on as many times as they liked) and answered the question "What characteristics come to mind when you hear this person speak?"[22] Listeners indicated where they thought the speaker was from and rated the speaker on two 6-point scales: *this speaker is definitely from California–this speaker is definitely not from California* and *this speaker sounds just like me–this speaker doesn't sound at all like me*. Listeners were also asked to give their reasoning for their answers to these questions ("What about their speech makes them sound like that?") to the best of their ability, as a way to assess the availability of listeners' attitudes toward speakers (see 1.4. Folk linguistics, above). After 15 randomly selected trials, listeners filled out a short demographic questionnaire.

Because I was more concerned with getting a large sample of data in a relatively short time than drawing from a regionally diverse sample of Californians, listeners were recruited through the UC Davis Psychology Research Participation System. This proved to be a highly efficient recruitment method, as 124 listeners took the survey in just 18 days.

3.3.4. Pretesting results

The pretesting data were analyzed to find the attributes that were most frequently mentioned by listeners. An initial list of over 180 potential attributes was generating through

[22] In the introduction to the task, listeners were given the examples of "intelligent," "rude," and "like a teenager."

listeners' responses, the attributes from the pilot study, and brainstorming. I counted the number of times an attribute appeared in a narrative answer and looked for clusters of related attributes.

This analysis resulted in the following seven scales for the main task: *feminine–masculine, Californian–not Californian, fast–slow, young–old, confident–not confident, relaxed–excited, friendly–not friendly*. Five additional scales were added to this list. The scales *familiar–unfamiliar* and *sounds like me–doesn't sound like me* were added to facilitate a comparison of listeners' attitudes with their beliefs about the speech of their communities and their own speech. The scale *sounds like a Valley girl–doesn't sound like a Valley girl* was added to assess the degree to which perceptions of 'Californian-ness' are tied up in the highly visible Valley girl stereotype, a label that itself surfaced numerous times in the pretesting results. The scales *rich–poor* and *suitable–not suitable* (for a job requiring speaking to an audience) were added to measure attitudes' status dimension. These scales also responded to previous research on the social meanings of the California Vowel Shift: D'Onofrio's (2014) finding that ambiguous *sock–sack* tokens were more likely to be labeled *sack* in a categorical perception task if the speaker was labeled as a "business professional," and Eckert's (2008b:29) suggestion that the CVS has been "gendered, raced, and classed" to index Whiteness, carefreeness, and privilege (with the *feminine–masculine* scale also measuring gendering and the *relaxed–excited* scale also measuring carefreeness). In addition, the scale *suitable–not suitable* reinforced the framing of the main task as an evaluation of candidates for a radio job (see 3.4.4. Main study survey, below).

As mentioned above, the secondary goal of the pretesting task was identifying excerpts or excerpt features that elicited unusually strong affective responses. While all excerpts were judged to be suitable in this respect, listeners attended at a particularly high rate to speakers' use of *like*, which is not surprising given the status of *like* as a stereotypical marker of California

English (Dailey-O'Cain 2000). *Like* tokens were edited out of five excerpts using the same procedure (and prioritization of naturalness) used for disfluencies (see 3.2.2. Conversation and task design, above). Syntactic considerations limited the number of tokens that could be eliminated; whereas focuser *like* ("the cat wanted to **like** play in the suitcase," Exc12) and approximator *like* ("he kinda **like** falls back," Exc5) can be removed without making the sentence ungrammatical, this is not the case for quotative *be like* ("Stu **was like** get out of my suitcase," Exc10) or conjunction *like* ("then he has an idea **like** I got to pack too," Exc16). Some tokens of focuser or approximator *like* could not be deleted without the excerpt sounding unnatural, so these tokens remained. No pretesting listeners reported that they suspected that some of the excerpts had been edited. Appendix D displays transcripts of the final edited versions of the excerpts that served as the basis for the creation of matched guises in the main study.

For the sake of evenness in the speaker sample (i.e., having one speaker in each of the 12 demographic cells), one of the two Southern California male Caucasian speakers, SC07-UK, was eliminated from the sample for the main study. Although both Caucasian male Southern California speakers received similar attributes in the pretesting data, I reasoned that this speaker could be more useful in recruiting his personal contacts for the main study, which would only be possible if his voice was not among the stimuli.

3.4. Main study

Stimuli were created by acoustically manipulating each excerpt to create a pair of guises: a California-shifted guise and a conservative (non-shifted) guise. In all, 91 TRAP tokens and 61 GOOSE tokens were manipulated twice (once for each guise) across the 24 excerpts. Listeners

took part in the dialect recognition task via an online study, and the results of this task were analyzed via Bayesian hierarchical modeling.

The following sections describe the methods for the main study: how this task fulfilled considerations of opacity, naturalness, and generalizability (3.4.1), the calculation of vowel manipulation targets (3.4.2), the acoustic manipulation process itself (3.4.3), the setup of the online survey (3.4.4), listener recruitment methods (3.4.5), and data analysis methods (3.4.6).

3.4.1. Matched-guise study design

In the Introduction chapter, I discuss three criteria for valid matched-guise tasks: opacity, naturalness, and generalizability (see 1.4.1. Language attitudes, above). These criteria led to my choosing to use spontaneous speech from the retells for stimuli. As mentioned above (see 3.2.1. Retell task design), the retell task limits content variation without limiting spontaneity, and the pretesting task ensured that content did not vary to an extent that it jeopardized generalizability. Because I was interested in perceptions of the intersection of California English features with region, gender, and ethnicity, I decided that excerpts from all stimulus speakers would be manipulated (i.e., there were no non-manipulated 'filler trials'). Opacity was satisfied in part by the online survey's framing as a task to "help us evaluate candidates for a local radio show job in California that involves story telling" (it was not stated who "us" was); this framing lessened the chance that listeners would become aware of the guise manipulation, as it encouraged listeners not to focus on minute phonetic details of the stimuli. In addition, no listeners heard more than one stimulus from any speaker, a design feature that supported opacity by preventing listeners from directly comparing two versions of the same excerpt (see 3.4.4. Main study survey, below).The nature of this study, in which matched guises were created from multiple speakers' spontaneous speech by using resynthesis to embed resynthesized vowels

within carrier phrases, presented unique challenges to satisfying opacity, naturalness, and generalizability; it is perhaps for this reason that, to my knowledge, no matched-guise studies have attempted this multi-speaker resynthesized-vowel design. I determined that acoustic manipulation was the only suitable method for creating the guises in this study. First, in general Californians apparently do not have an available notion of variation in TRAP backing and GOOSE fronting, so it would be difficult to get most speakers to perform the distinction. Second, even if I found a sufficient number of speakers who could perform these distinctions, the continuous nature of vowel variation (unlike categorical variables like (ing) or *like*) would make it impossible to guarantee generalizability across speakers; there would be no guarantee, for example, that multiple speakers' performances of TRAP backing would be equally exemplary of a California-shifted TRAP vowel. Resynthesis, by contrast, necessitates the specification of target manipulation values, which allows the researcher to fine-tune guises to satisfy generalizability. Finally, the flexibility of acoustic manipulation meant that the stimuli could be drawn from more spontaneous speech rather than relying on speech in a more formal style.

3.4.2. Manipulation target calculation

Once I determined that acoustic manipulation would be used to create guises, I had to find manipulation *targets*, the formant values that would represent conservative vs. California-shifted TRAP and GOOSE tokens. Figure 3.7 represents these guises in schematized fashion. As I detail more specifically below, I calculated targets for each speaker and vowel based on the production of that vowel by both the individual speaker and the overall group.

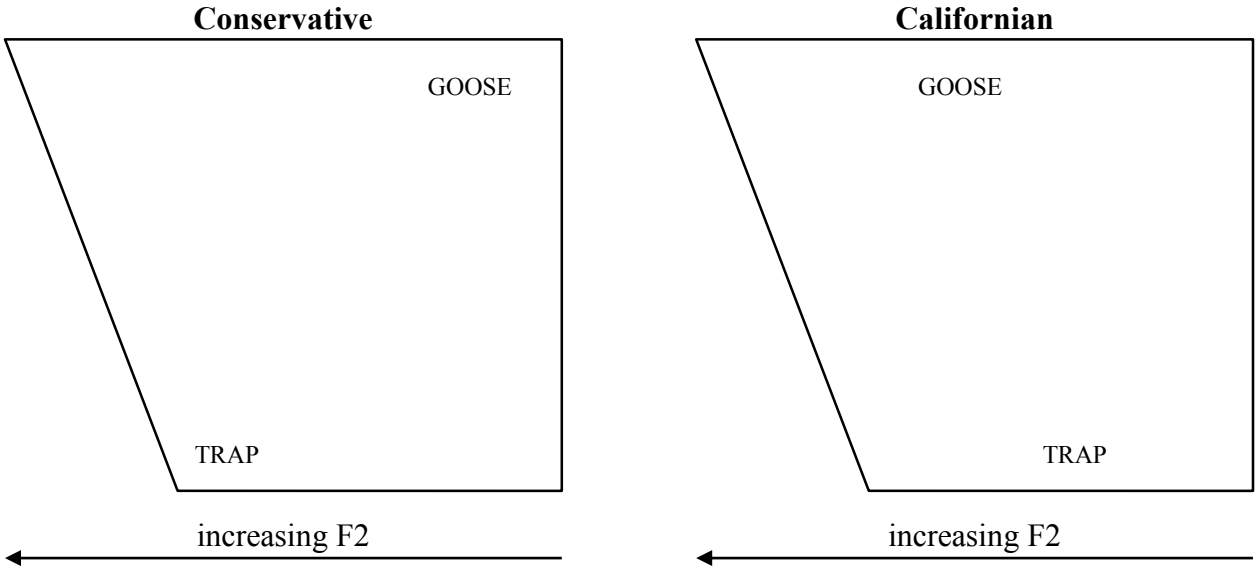


Figure 3.7. Schematic of conservative vs. Californian guises.
The conservative guise featured relatively front (higher F2) TRAP and relatively back (lower F2) GOOSE, whereas the Californian guise featured relatively back (lower F2) TRAP and relatively front (higher F2) GOOSE.

This process presented two main challenges for fulfilling generalizability and naturalness. First, individual speakers vary considerably in the size of their vowel spaces as a function of variation in vocal tract size, with females typically having larger vowel spaces than males (Johnson 2003). As a result, a given feature's manipulation target cannot be defined absolutely (i.e., as a single formant value) but instead must be defined relative to individual speakers' vowel spaces. For example, as Figure 3.8 below shows, while a token of GOOSE with an F2 value of 2100 Hz would sound rather backed for CV03-NG, a female speaker whose average GOOSE F2 is 2273 Hz, it would also sound very fronted for CV05-FC, a male speaker whose average FLEECE F2 is 1933 Hz and average GOOSE F2 is 1575 Hz.[23] This idea, which extends back to Joos's (1948) original study of acoustic phonetics, underlies the basic motivation for the common sociophonetic practice of speaker normalization (Watt et al. 2011).

[23] Vowel plot figures in this chapter (Figures 3.8, 3.9, 3.10, 3.13, and 3.14) were produced using the vowels R package (Kendall & Thomas 2010).

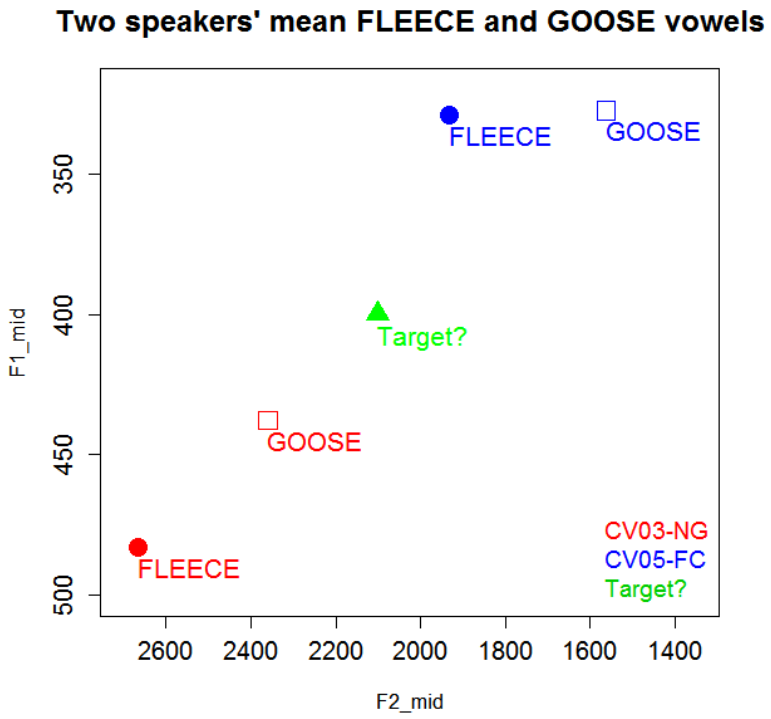


Figure 3.8. Average FLEECE and GOOSE vowels for CV03-NG and CV05-FC, with a hypothetical F2 target of 2100 Hz.
CV03-NG is a female speaker and thus has a larger vowel space than CV05-FC, a male speaker.

Moreover, the continuous nature of vowel variation presents a challenge for fulfilling generalizability. Even if every speaker had an equally sized vowel space, it would still be impossible to identify, a priori, a target that exemplified a certain feature to a greater extent than any other possible target (e.g., it is impossible to evaluate whether the prototypically fronted GOOSE has an F2 of 2250 Hz, 2220 Hz, 2286 Hz, 2194 Hz, etc.). As a result, the target calculation procedure used in this study relied on speakers' natural range of vowel variation to adequately represent extreme values among this group of speakers.

In the context of this study, naturalness meant that all manipulated tokens had to be within a plausible range of productions for each speaker; in addition, the stimuli could not sound computer-edited to non-linguist listeners (who were not made aware of the computer

manipulation in this task). Generalizability meant that the Californian guises had to sound equally California-shifted, and the conservative guises equally non-shifted. In other words, when listeners reacted to the Californian (or conservative) tokens of any one speaker, that had to be qualitatively the same thing as when listeners reacted to the Californian (or conservative) tokens of any other speaker.

In order to fulfill both naturalness and generalizability, my approach for calculating targets balanced individual speakers' vowel production with that of the overall group. In short, manipulation targets for a given speaker and vowel were defined as the average between the speaker's range of productions of the vowel and other speakers' range of productions.[24] This approach thus dealt with the dual issues of interspeaker vowel space variability on the one hand and the lack of prototypically fronted or backed targets on the other by generating tokens that had their basis in individual speakers' actual productions. Specifically, the target calculation had the following steps:

1. Each speaker's range of variation in TRAP F2 was calculated as the difference between their 2nd-highest F2 and 2nd-lowest F2 for all TRAP tokens in their retells (Table 3.3); I used 2nd-highest and 2nd-lowest values because several speakers' highest and/or lowest F2 values were outliers. For example, SC02-KH's 2nd-lowest TRAP F2 was 1433 Hz and his 2nd-highest TRAP F2 was 1671 Hz, giving him a TRAP F2 range of 238 Hz. Speakers' ranges of variation were also

[24] This process was the result of some trial-and-error to find targets that satisfactorily balanced naturalness and generalizability for all speakers. (This is another benefit of using vowel resynthesis for matched-guise studies: the researcher can experiment with different manipulation targets on their own time rather than having speakers produce all possible targets.) Among the procedures that failed to produce satisfactory manipulated tokens were using the same targets for all speakers of the same gender, using targets based only on individual speakers' ranges of production, and using targets derived from benchmarks defined by previous studies (i.e., Table 3.2).

calculated separately for GOOSE F2 using the same procedure as for TRAP F2. For example, SC10-NM's 2nd-lowest GOOSE F2 was 1370 Hz and her 2nd-highest GOOSE F2 was 2200 Hz, giving her a GOOSE F2 range of 830 Hz.

		TRAP			GOOSE		
Speaker code	**Gender**	***n***	**Min–Max**	**Range**	***n***	**Min–Max**	**Range**
BA07-SQ	Female	23	1752–2004	252	10	2055–2344	289
BA11-DM	Female	33	1526–1761	235	22	1573–2133	560
BA05-TN	Male	43	1513–1714	201	18	1307–1784	476
BA01-CH	Male	76	1474–1755	281	37	1245–1847	602
CV03-NG	Female	28	1673–1990	317	9	1948–2502	554
CV01-NK	Female	25	1647–1862	215	13	1665–2268	603
CV05-FC	Male	16	1390–1762	373	6	1515–1655	140
CV04-CX	Male	57	1490–1809	319	19	1496–2077	581
SC06-KY	Female	29	1699–1932	233	18	1229–1864	635
SC10-NM	Female	33	1462–1821	359	14	1370–2200	830
SC02-KH	Male	23	1433–1671	238	16	1313–1927	614
SC05-KB	Male	52	1486–1877	390	33	1140–2071	931
Averages	**Female**	**29**	**1626–1895**	**269**	**14**	**1640–2219**	**578**
	Male	**45**	**1464–1765**	**300**	**22**	**1336–1893**	**557**

Table 3.3. TRAP and GOOSE tokens in main study stimulus speakers' retells.
All values in Hz. *n* indicates token count, Min indicates 2nd-lowest F2, Max indicates 2nd-highest F2, and Range indicates F2 range.

2. The averages of speakers' 2nd-highest and 2nd-lowest F2 were each calculated by gender and vowel, giving four average ranges (female TRAP F2, male TRAP F2, female GOOSE F2, male GOOSE F2). Female speakers' average TRAP F2 spanned 1626–1895 Hz (a range of 269 Hz), for example. These average ranges are displayed in the bottom two rows of Table 3.3.
3. For each female speaker, conservative TRAP F2 targets were calculated by averaging the speaker's 2nd-highest TRAP F2 and the average of female speakers' 2nd-highest TRAP F2. For example, CV01-NK's 2nd-highest TRAP F2 was 1862 Hz and the average of female speakers' 2nd-highest TRAP F2 was 1895 Hz, giving

CV01-NK a conservative TRAP F2 target of 1879 Hz. Similarly, conservative TRAP F2 targets were calculated for each male speaker by averaging each male speaker's 2nd-highest TRAP F2 and the average of male speakers' 2nd-highest TRAP F2.

4. Conservative GOOSE F2 targets were calculated for each female and male speaker via a similar process to conservative TRAP F2 targets, instead using the 2nd-*lowest* GOOSE F2 as the basis for targets.
5. Californian TRAP F2 targets were calculated for each female speaker by subtracting the average range of female speakers' TRAP F2 production (269 Hz) from the speaker's conservative TRAP F2 target. For example, CV01-NK's conservative TRAP F2 target was 1879 Hz, so her Californian TRAP F2 target was 1610 Hz. Similarly, Californian TRAP F2 targets were calculated for each male speaker by subtracting the average range of male speakers' TRAP F2 production (300 Hz) from each male speaker's conservative TRAP F2 target.
6. Californian GOOSE F2 targets were calculated for each female and male speaker similarly, but by *adding* the average range of females' or males' GOOSE F2 production *minus* 100 Hz (478 Hz for females, 457 Hz for males). For example, SC02-KH's conservative GOOSE F2 target was 1325 Hz, so his Californian GOOSE F2 target was 1782 Hz. This 100-Hz adjustment was implemented because prior to making this adjustment, some Californian GOOSE tokens were so fronted that they violated naturalness and/or were computationally impossible to implement (e.g., without this adjustment, BA07-SQ's Californian GOOSE F2 target would have been 2426 Hz—rather close to her average FLEECE F2 of 2520 Hz). I

determined through auditory inspection that manipulated tokens based on this adjusted target were more than adequate to represent fronted GOOSE.

7. Given that GOOSE fronting is promoted by preceding coronal segments, GOOSE F2 targets were adjusted for the four GOOSE tokens that did not follow coronals: *removed*, *food*, and two tokens of *room*. Labov's (2010:264) multivariate analysis of GOOSE data from the Atlas of North American English showed that the postcoronal context adds 244 Hz to the expected F2 of GOOSE.[25] As a result, both the conservative and Californian targets for two postlabial tokens, *removed* and *food*, were adjusted by −244 Hz. Labov's analysis does not specify a value for the effect of preceding /r/ on GOOSE, but his graph of frequent words appears to indicate that preceding /r/ has a medium-sized effect on GOOSE fronting (267), so the targets for two tokens of *room* were adjusted by −122 Hz.[26] A post-/j/ token, *unusually*, was deemed to sound natural using the general (i.e., postcoronal) GOOSE targets, so the targets for this token were not adjusted.[27]
8. For the majority of tokens, only F2 was manipulated, but initial runs of the manipulation procedure revealed nine tokens for which F3 also needed to be manipulated to preserve naturalness; F1 was not manipulated for any tokens. For eight GOOSE tokens, F3 was low enough that when F2 was raised in the initial versions of the Californian guise, F2 and F3 overlapped in a way that produced

[25] My analysis of Cory Holland's (2014) California English vowel data revealed that postcoronal GOOSE was on average 222 Hz further front than non-postcoronal GOOSE, a strikingly similar result to Labov's 244 Hz result, which is based on speakers from across North America. This similarity across speech communities suggests a 'pure' phonetic conditioning for postcoronal GOOSE fronting.
[26] Without the adjustment, the California-shifted *room* tokens were so fronted that they sounded like *rim*; with the adjustment, they simply sounded like fronted *room*.
[27] The prelateral environment is known to inhibit GOOSE fronting (Holland 2014; Labov 2010), but as no GOOSE tokens in the excerpts were prelateral, no adjustments were made for prelateral targets. No environment-based adjustments were made to TRAP targets.

non-speech-like chirping sounds. In order to avoid these problematic overlaps, F3 was set to be raised by 100, 200, or 300 Hz (depending on the token) in only the Californian guise and for only these eight tokens. After this adjustment, the chirping sounds were eliminated and there was no appreciable difference between the quality of Californian GOOSE tokens with raised F3 or non-raised F3. In addition, one TRAP token sounded highly nasalized in its conservative guise (but not in its Californian guise), so F3 was set to be raised by 100 Hz in only the conservative guise; after this adjustment, the percept of nasality was mitigated. For all nine of these tokens, F3 was manipulated prior to F2.

Table 3.4 displays the outcome of the target calculation procedure: each speaker's TRAP F2 and GOOSE F2 targets for the conservative and Californian guises. The plots of conservative and Californian targets (Figures 3.9 and 3.10) demonstrate that these targets are in accord with the schematized guises in Figure 3.7. For conservative targets GOOSE is further back (lower in F2) than TRAP, whereas for Californian targets TRAP is further back than GOOSE—importantly, this relationship is true for all speakers.

		TRAP		GOOSE	
Speaker code	**Gender**	**Cons**	**Cali**	**Cons**	**Cali**
BA07-SQ	Female	1950	1681	1848	2326
BA11-DM	Female	1828	1559	1607	2085
BA05-TN	Male	1739	1439	1322	1779
BA01-CH	Male	1760	1459	1291	1748
CV03-NG	Female	1943	1674	1794	2273
CV01-NK	Female	1879	1610	1653	2131
CV05-FC	Male	1764	1463	1425	1883
CV04-CX	Male	1787	1486	1416	1873
SC06-KY	Female	1913	1645	1435	1913
SC10-NM	Female	1858	1589	1505	1984
SC02-KH	Male	1718	1418	1325	1782
SC05-KB	Male	1821	1520	1238	1695

Table 3.4. F2 manipulation targets for each main study stimulus speaker. All values in Hz. Targets are Cons[ervative] or Cali[fornian].

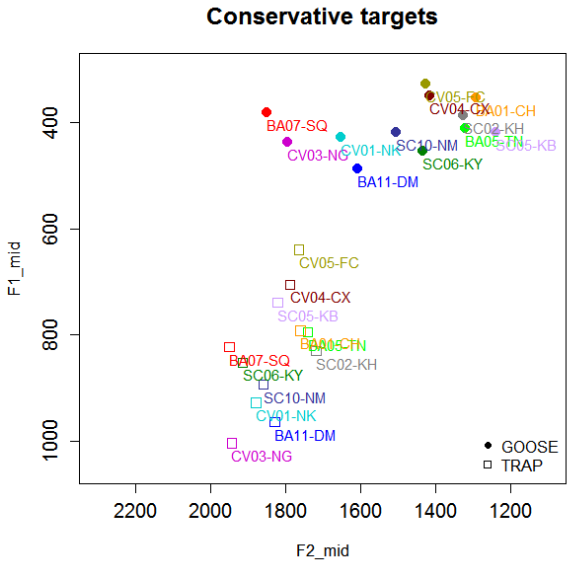


Figure 3.9. Stimulus speakers' conservative targets. Note that targets' F1 values represent each speaker's average F1 for that vowel; F1 was not manipulated.

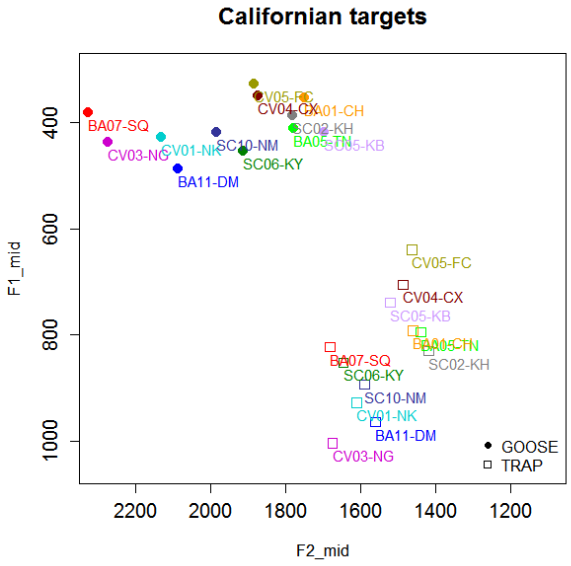


Figure 3.10. Stimulus speakers' Californian targets. Note that targets' F1 values represent each speaker's average F1 for that vowel; F1 was not manipulated.

3.4.3. Implementation of manipulation

Underlying vowel resynthesis methods is the source–filter theory of speech production, which posits that speech sounds originate from a glottal source (i.e., the vibration of vocal folds)

and pass through a filter defined by the positions of articulators of the vocal tract (Johnson 2003). For American English vowels, the filter is defined by the position of the tongue and the rounding of the lips, and these articulators affect the frequencies at which the sound from the glottal source is amplified or dampened (i.e., formant structure). Thus, whereas characteristics like phonation type, pitch, and loudness are properties of the source, the acoustic characteristics that most strongly correlate with the perception of vowel location, formant frequencies, are properties of the filter (Styler 2015).

Vowel resynthesis utilizes source–filter theory (and in particular, the assumption that source and filter are relatively independent) by performing linear predictive coding on a vowel to estimate the acoustic effects of the filter (formants), inverse-filtering the vowel to derive the underlying source, modifying the filter, and passing the source through the modified filter. In essence, these steps comprise the basic procedure that was used to create stimuli for the main study. In order to preserve naturalness (and to automate the manipulation process via Praat script), however, I had to make several modifications in order to ensure that this basic procedure worked for the 91 TRAP tokens and 61 GOOSE tokens manipulated in the 24 excerpts. Appendix E gives step-by-step details on how the manipulation process was carried out.

One advantage of using an automated procedure is that the same manipulation steps applied to every token. However, I had to replace five tokens (one TRAP, four GOOSE) in the original excerpts because of difficulties manipulating the original tokens. Because the problematic tokens were replaced in the original excerpts prior to manipulation, the replaced tokens were manipulated for both guises. Four GOOSE tokens simply resisted being manipulated to either or both of their targets. As a result, these four tokens were replaced by different tokens from elsewhere in the same speaker's retell (but not the same excerpt). For some of these vowels,

the source from the original token was retained and passed through a filter (formant structure) derived from a different token; for others, the original, problematic token was replaced completely by splicing in a new token (both source and filter). The pitch and intensity of the new token were adjusted in order to make the new token sound as natural as possible within the larger phrase. In addition, in one excerpt, a speaker's pronunciation of the vowel in *bag* was raised to [e:], a pattern that is typical of the Pacific Northwest (Wassink 2015) but has not been reported in California. In order to avoid this vowel overly influencing listeners' perception of the speaker's origin, this vowel's formants were replaced with that of a different prevelar TRAP token (*back*) from elsewhere in the same speaker's retell.

After the acoustic manipulation process was run for all 24 excerpts, a trained phonetician listened to the manipulated stimuli to gauge naturalness and generalizability, which led to several small adjustments (e.g., re-splicing a vowel). Stimuli were deemed to be satisfactory after these adjustments.

Figures 3.11 and 3.12 demonstrate the end result of this process for TRAP and GOOSE tokens, with original versions of these tokens compared to the same tokens in the conservative and Californian guises.

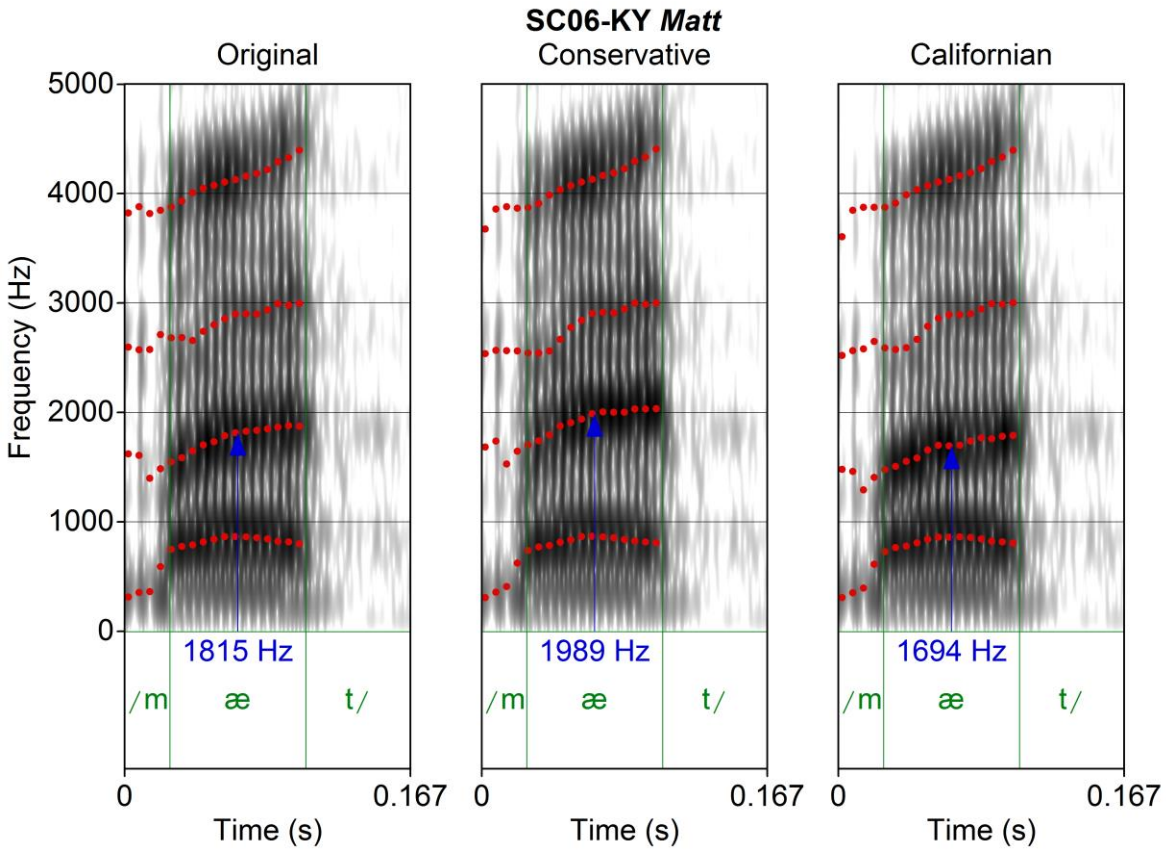


Figure 3.11. Spectrograms and formant tracks of original, conservative, and Californian versions of the token *Matt* by SC06-KY.
Formant tracks are in red. Blue arrows and text indicate F2 value at the vowel midpoint.

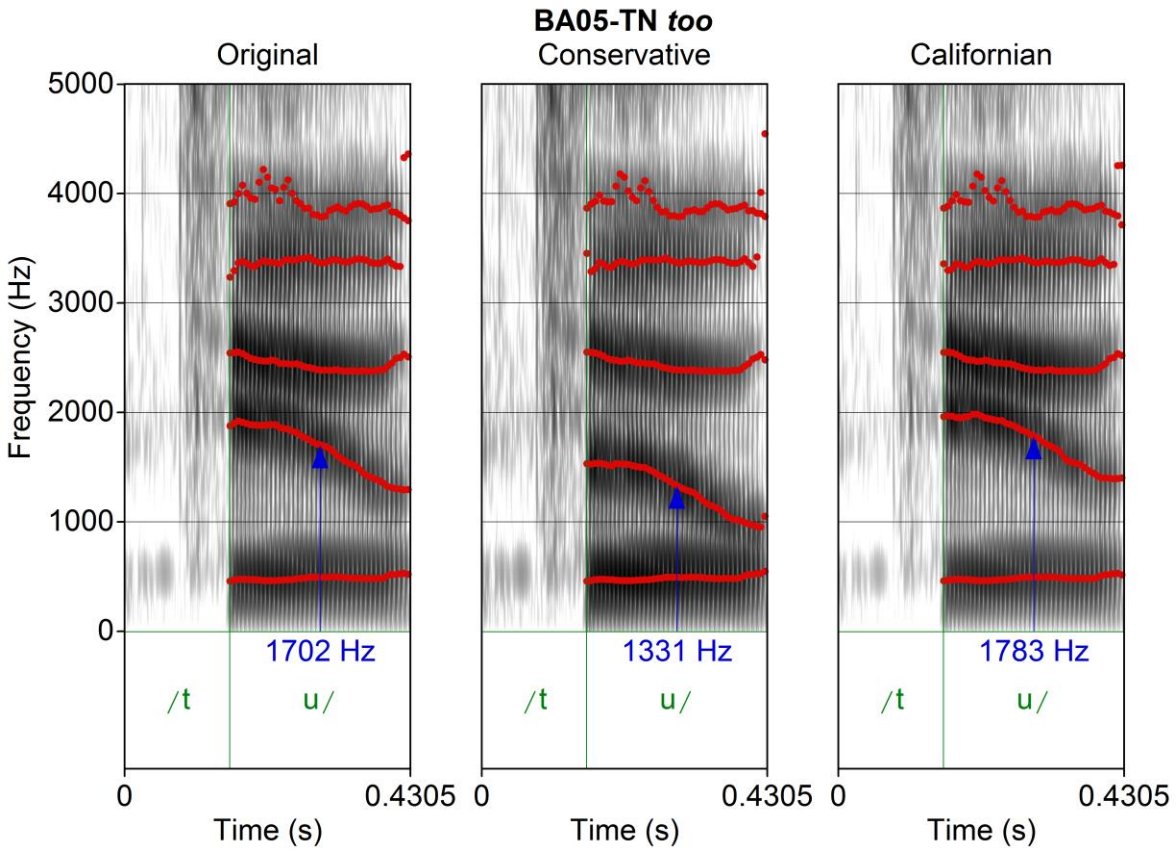


Figure 3.12. Spectrograms and formant tracks of original, conservative, and Californian versions of the token *too* by BA05-TN.
Formant tracks are in red. Blue arrows and text indicate F2 value at the vowel midpoint.

This manipulation procedure succeeded in creating manipulated tokens that were acceptably close to F2 target values. I defined a token to be "acceptably close" to the target if the token–target F2 difference was less than 1 just noticeable difference (JND), the minimum difference that human perceivers can detect between two stimuli (in this case, the minimum detectable difference in formant frequency); for TRAP F2, 1 JND is 33.09 Hz, and for GOOSE F2, 1 JND is 21.86 Hz (Kewley-Port & Watson 1994:492). On average, manipulated TRAP tokens were 11.20 Hz off target in F2 and manipulated GOOSE tokens were 8.31 Hz off target in F2, both

well under 1 JND.[28] All 91 TRAP tokens were within 1 JND of the target for both guises; 3 of 61 conservative GOOSE tokens and 4 of 61 Californian GOOSE tokens were more than 1 JND off target.[29] In Figures 3.13 and 3.14, each speaker's original and manipulated tokens are compared with their conservative and Californian targets. It is clear from these graphs that while there is some variability in the manipulated tokens' F2, in general they are close to their targets irrespective of the location of the original tokens.

[28] For TRAP, conservative tokens (6.58 Hz off target) were on average more accurate than Californian tokens (15.82 Hz off target). For GOOSE, Californian tokens (5.48 Hz off target) were on average more accurate than conservative tokens (11.15 Hz off target).

[29] It is unclear to me why these tokens ended up more than 1 JND off target, as the manipulation procedure should have continued to shrink the distance between the token and target value until it was less than 1 JND (Appendix E). On average, these tokens were 25.87 Hz off target, which is not substantially larger than 1 JND (21.86 Hz); indeed, auditory inspection of these off-target tokens revealed that they sounded adequately similar to other manipulated tokens that generalizability was not at risk.

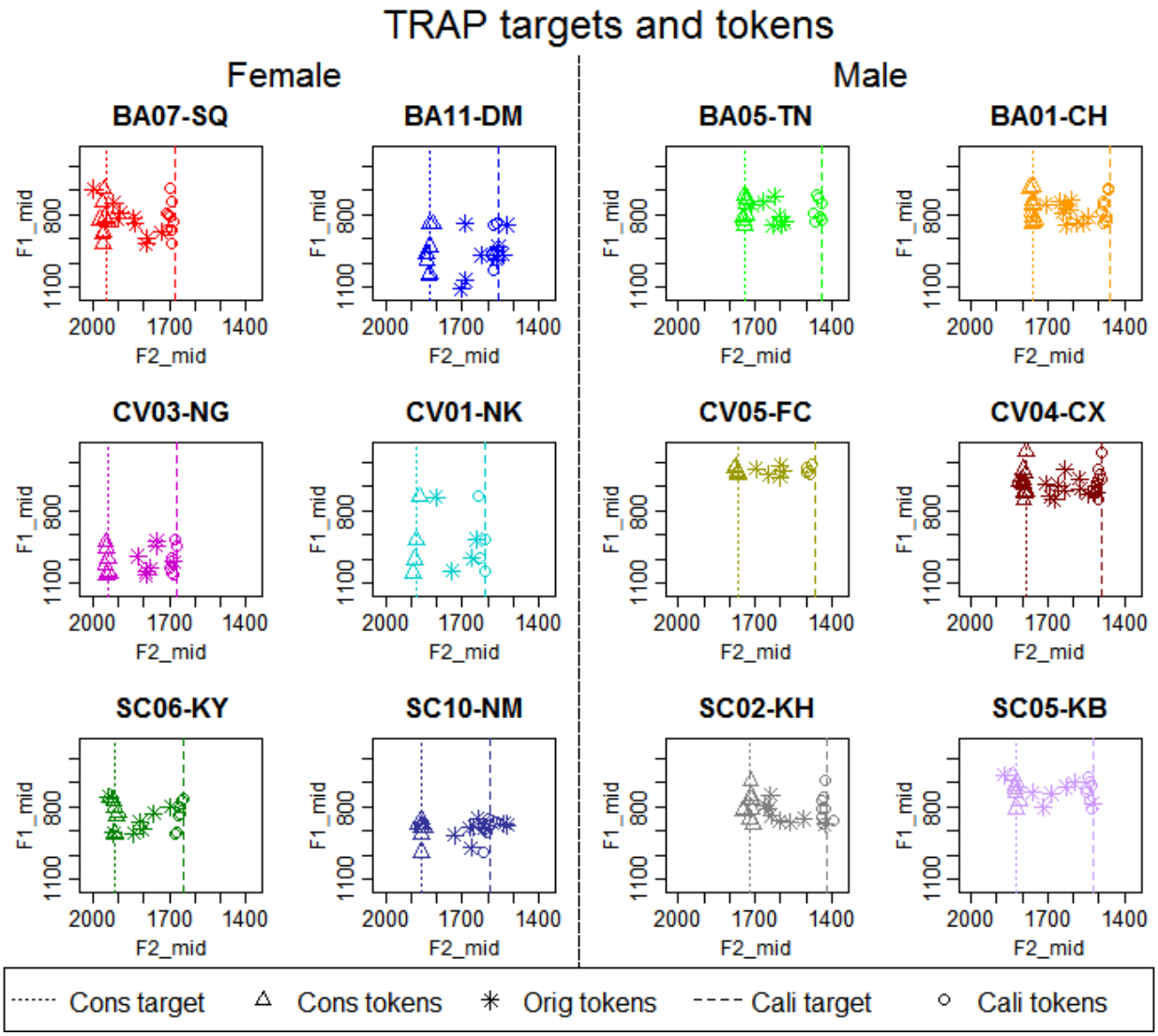


Figure 3.13. TRAP targets and original and manipulated TRAP tokens for each main study stimulus speaker. Targets are denoted by vertical lines since targets were only defined for F2, not F1; tokens are denoted by symbols.

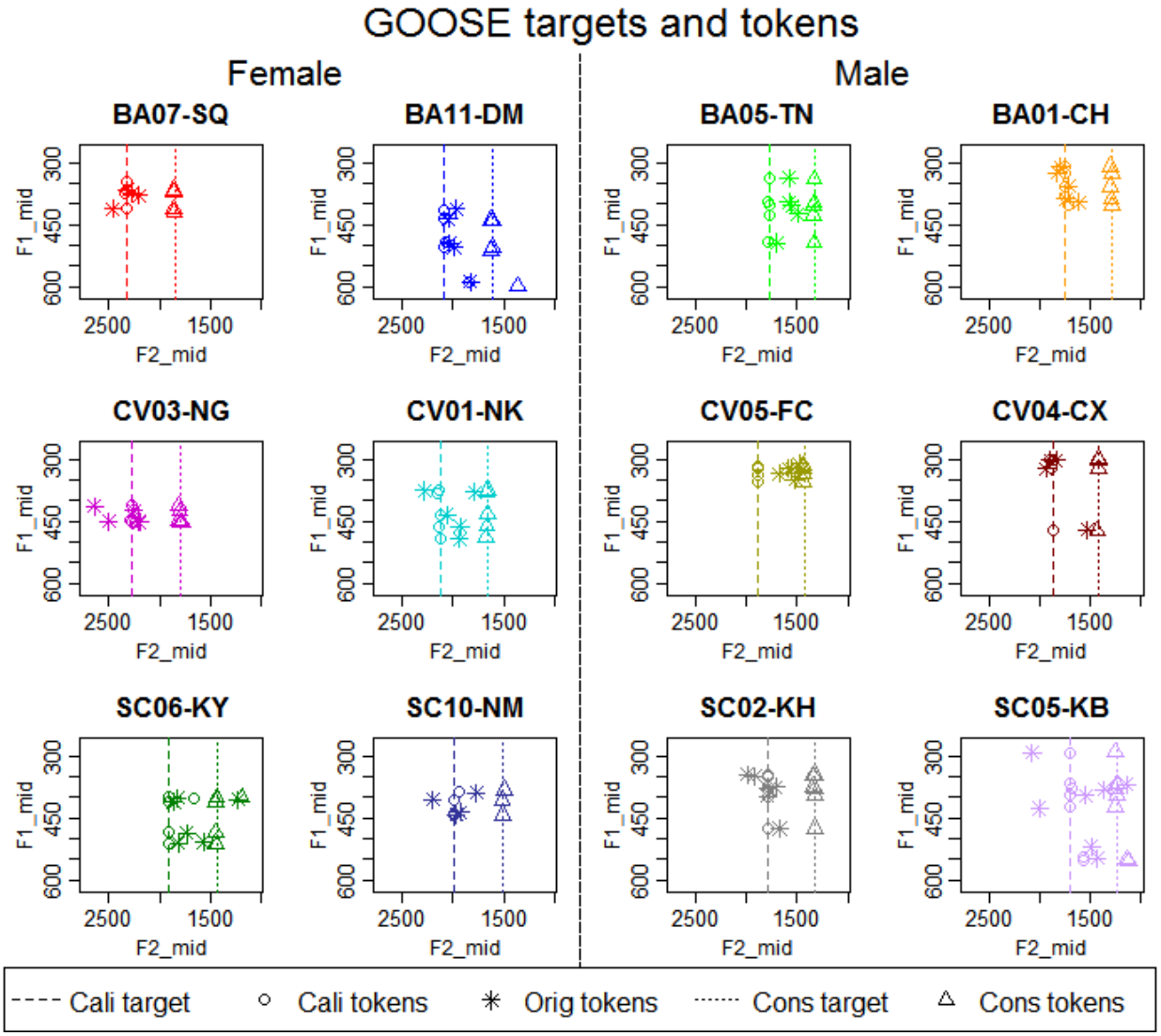


Figure 3.14. GOOSE targets and original and manipulated GOOSE tokens for each main study stimulus speaker. Targets are denoted by vertical lines since targets were only defined for F2, not F1; tokens are denoted by symbols. Four manipulated tokens (one for BA11-DM, one for SC06-KY, and two for SC05-KB) appear substantially offset from the targets. These tokens were non-postcoronal tokens, and as such both their conservative and Californian targets were adjusted to account for the fact that the postcoronal environment promotes GOOSE fronting (see 3.4.2. Manipulation target calculation, above); that is, these tokens are not actually off target, as their targets are different from the majority of tokens'.

Figures 3.13 and 3.14 reveal an unintended side effect of the manipulation process: F1 changed for many tokens, in some cases dramatically, even though F1 was not manipulated. For example, BA01-CH's lowest original TRAP F1 is 744 Hz, but two of his manipulated TRAP tokens are under 700 Hz for both guises. On average, F1 changed by 16.05 Hz for TRAP tokens and 3.24 Hz for GOOSE tokens. In general, F1 changed to a similar degree for both guises of the same

token (though this was not the case for all tokens, as with BA07-SQ's stray raised Californian GOOSE token). On average, the difference between the same token's conservative and Californian F1 was 5.46 Hz for TRAP and 3.25 Hz for GOOSE, well below Kewley-Port and Watson's (1994:492) JND values of 19.55 Hz for LOT F1 (TRAP F1 is not given) and 10.45 Hz for GOOSE F1. F3 values were slightly more off target, as the average difference between the same token's conservative and Californian F3 was 12.15 Hz for TRAP and 14.16 Hz for GOOSE (not counting tokens for which F3 was intentionally raised). These changes in F1 and F3 were most likely a consequence of inherent difficulties in completely decoupling the filter from the source (see fn. 5 of Appendix E).

In addition to its accuracy, this method also proved to be relatively speedy. In total, manipulating 152 tokens twice across 24 excerpts took slightly under three minutes, about a second per token. Among these 304 manipulations, the median number of iterations of the main manipulation step (see Appendix E) was seven, although it was not uncommon for some tokens to undergo 20 or more iterations.

3.4.4. Main study survey design

The main study survey was an online survey hosted on Qualtrics (like the pretesting survey). I opted for an online survey because it could be easily deployed to all parts of the state and shared via email and social media, as well as because of the randomization and data collection tools available on Qualtrics. When listeners started the survey, after agreeing to a consent, they were informed that they would "help us evaluate candidates for a local radio show job in California that involves story telling" (it was not stated who "us" was). This framing, similar to that used by Labov et al. (2006a), was included so listeners would not shy away from sharing their attitudes (especially negative attitudes). This framing also provided context for

otherwise decontextualized questions about a speaker's personal characteristics, as well as external motivation for listeners to complete the task honestly (i.e., aiding in quality control of radio hosts): "We are trying to get a sense of how listeners might react to each of these job candidates, so there are no right or wrong answers!" I included the "story telling" aspect of the framing, explicitly mentioning the videos with Matt the cat and Stu, to alert listeners that they would be hearing highly similar content. Listeners were debriefed about the framing at the end of the survey.

The main part of the survey consisted of six randomly selected listening trials, each from a different speaker. For each listener, Qualtrics randomly chose six of the twelve speakers, then chose one of the four possible trials from each speaker; as a result, no listeners heard more than one trial from any speaker. This arrangement is schematized in Figure 3.15. The survey did not include a mechanism to ensure that listeners heard both conservative and Californian stimuli in their six trials, but because of the survey design listeners had a 1-in-64 chance of only hearing one type of guise or the other.[30]

[30] The chances that listeners heard only one of two genders or one of two ethnicities were 1 in 924. Because listeners heard six speakers and there were four speakers from each region, listeners could not hear only one region; the chances that listeners heard only two of three regions were 1 in 33.

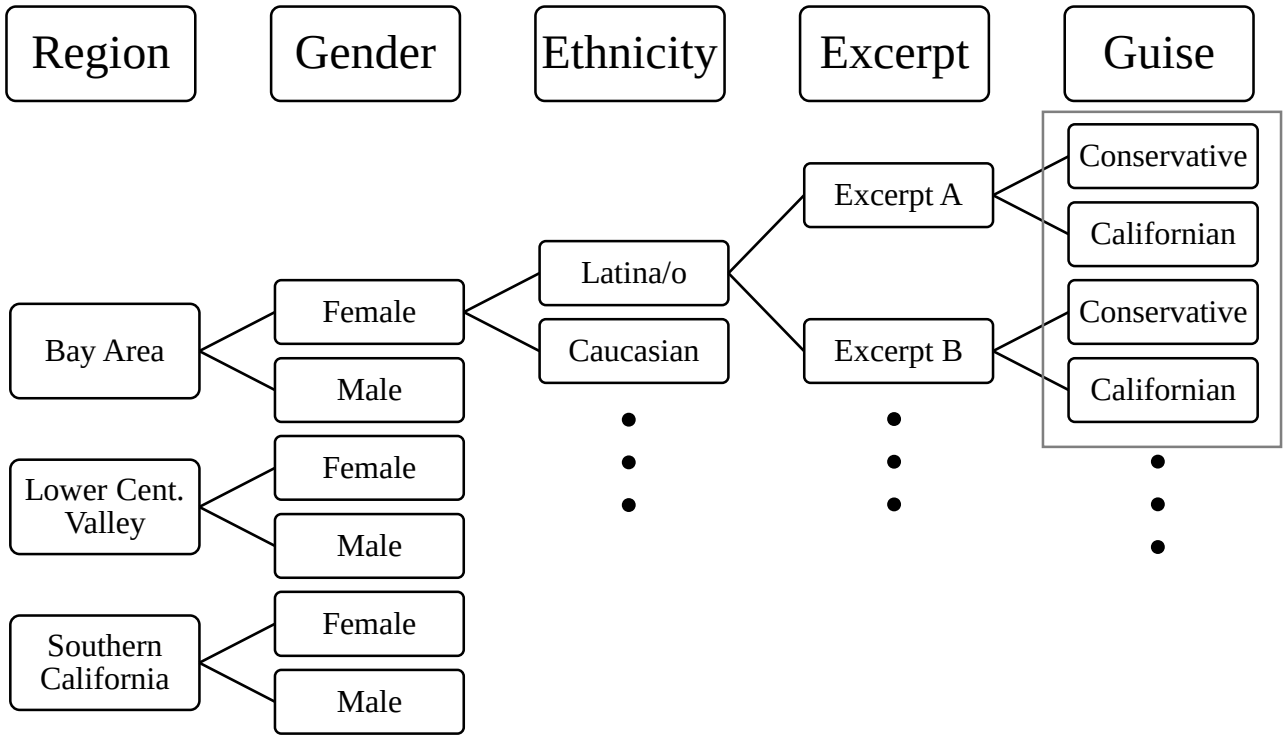


Figure 3.15. Schematic of trials in the main task.
Ellipses indicate that the boxes in each column continue beyond what is shown. The gray box surrounds a single grouping of four trials from a given speaker: the conservative and Californian guises of this speaker's Excerpt A, and the conservative and Californian guises of this speaker's Excerpt B. There were 12 of these trial groupings in total, one for each speaker.

In each trial, listeners heard a clip (they had the option to click on it as many times as they liked) and were asked to identify the speaker's region and rate the speaker on the 12 scales generated by the pretesting task (see 3.3.4. Pretesting results, above). For the regional identification question, speakers were given the choice of the Bay Area, Lower Central Valley, Southern California, or outside California (with a blank to specify this answer). This question was accompanied by a map of the state with each region outlined (Appendix F), the same map used for the pilot study without Far Northern California. The first scale question was always "How suitable do you think this speaker would be for a job requiring speaking to an audience?" with the poles *not at all suitable–highly suitable*. This scale was followed by "Please rate the speaker on the following scales" and the other 11 scales in an order that was randomized by trial.

Listeners responded to each scale by dragging a handle to a point on a continuous slider bar, as in Figure 3.16.[31] After both the regional identification question and the scales, listeners were asked, "What gave you that impression? (Be as specific about particular speech features as possible)." Listeners were also asked how many times they listened to the clip, as well as whether they recognized the speaker.[32]

Figure 3.16. Three slider bars from the scale rating task.

After completing six trials, listeners completed a demographic questionnaire (Appendix G). At the end of the questionnaire, listeners were given the option to upload a clip of themselves

[31] For the purpose of data analysis, Qualtrics translated this position into an integer between 0 and 100.
[32] This question, "Do you recognize this speaker as someone you know?", was included because listeners and speakers were recruited through similar channels (e.g., linguistics professors at certain universities), so it was possible that a listener may have been able to identify a certain speaker, meaning that trial would need to be excluded from the analysis. Many listeners seemed to interpret this question as a hypothetical (e.g., "I recognize this speaker as someone I would know," "yes, sounds like someone from NPR") rather than asking about specific speakers or simply answered "yes" with no specifics. No listeners correctly identified the specific speaker of the stimulus by name.

reading the modified version of "Comma gets a cure" that was used for interviews (Appendix B) so I could compare listeners' production and perception; few listeners opted to do so, however.

3.4.5. Listeners

The survey was active from late April to early June 2015. Listeners were recruited via "friend-of-friend" sampling through a variety of contacts: linguistics professors at universities in the Bay Area, Lower Central Valley, and Southern California (several of whom had aided in recruiting speakers for the interview stage), friends, colleagues, former students, and even interviewees with whom I felt I had established a friendly rapport (and who were not in the stimulus pool). To incentivize participation, I offered to randomly select three listeners to win $50 Amazon.com gift cards, and several professors offered extra credit.

The data set for the main study analysis is drawn from 97 listeners' responses to the online survey. One hundred four listeners completed the survey, but seven were removed from the data set because they were not from California (defined as identifying a hometown outside the state), did not self-identify as Californians, and/or knew too much about the research to be included in the data set. Each listener responded to six trials, so the total number of trials analyzed was 580 (two listeners experienced audio failure on one trial).

Listeners' home regions and mobility were determined via their responses to the "Please list all the places you have lived" and "What do you consider your 'hometown'?" questionnaire items.[33] Eighty-three listeners had hometowns in the Bay Area (26), Lower Central Valley (31), or Southern California (26). Fourteen listeners had hometowns in other parts of the state: Far Northern California (2), Greater Sacramento (9), and the Central Coast (3). Listeners were

[33] Including these items in tandem proved useful, as some listeners interpreted the latter question as a semantic or philosophical question (e.g., one listener answered, "where you lived the longest").

generally somewhat mobile; 18 reported having lived in the same city their entire lives, 11 in different cities in the same region, 37 in multiple regions of California, 4 in multiple Western states, 13 in multiple regions of the US, and 11 in multiple countries (mobility could not be determined for 3 listeners). Differences in mobility were not represented evenly across listener regions; just 3 of 26 Bay Area listeners and 6 of 26 Southern California listeners had lived only in their home region, as compared with 16 of 31 Lower Central Valley listeners (13 of whom had lived in the same city their entire lives).

The majority of listeners (65) self-identified as female, with 31 listeners self-identifying as male and 1 self-identifying as genderqueer. A narrow majority of listeners (49) identified as Caucasian American, 20 as Latina/o, 15 as Asian American, 10 as multiple ethnicities, and 2 as African American (1 listener declined to state an ethnicity). The listener sample skewed young, with a median age of 22, three quarters of listeners 26 or younger, and just nine listeners 40 or older. Most listeners reported speaking only English (27) or one other language (41).

The majority of listeners (70) were students (as evident from self-reported occupation and/or referral by a current professor). The majority of listeners (69) reported having studied linguistics; the questionnaire did not ask listeners to specify their level of experience, but 12 listeners reported having taken (or currently taking) a single linguistics class, 2 reported having taken multiple classes, and 8 reported being linguistics majors. This represents an over-representation of linguistic experience relative to the general population; it is not surprising that linguistics was over-represented in this survey given that its recruitment methods leaned on recruitment in linguistics classes. Indeed, 84.3% of students in the survey had had some linguistics experience, as compared with 37.0% of non-students.

3.4.6. Data analysis

The survey yielded two main types of data: regional identification data (map question) and language attitudes data (affective scales).[34] The analysis of this data set was carried out via Bayesian inference, which is based on the computation of distributions that assign different levels of credibility to possible values of given parameters (Kruschke 2015). Bayesian inference has several advantages over the methods more commonly used to conduct inferential statistics: frequentist inference. Frequentist methods assess the credibility of a hypothesized parameter value (null hypothesis) by computing the probability of observed data given the null hypothesis and a sampling distribution (*p* value). Because data are samples drawn from a wider population, a small proportion of data samples will be so far removed from the null hypothesis that they cast doubt on the credibility of the null hypothesis, even if the wider population is well described by the null hypothesis (and even in absence of systematic sampling bias). Frequentist inference thus defines a cutoff for null hypothesis credibility (a significance level, typically .05) below which analysts can be reasonably certain of the null hypothesis's falseness without overly raising the risk of false positives.

While Bayesian statisticians have raised numerous issues with this approach, the most important one for the purposes of the present study is multiple comparisons. In effect, when a frequentist analyst chooses a significance level, they are choosing the rate at which false positives are acceptable and expected; with a significance level of .05, for example, the analyst makes the assumption that 1 in 20 significance tests will yield a false positive (i.e., the rejection of a null hypothesis that actually describes a population well). This study tests the influence of

[34] The survey also yielded qualitative responses to the "What gave you that impression?" questions, but the analysis of these responses will be left to a future study. Listeners' production will also not be included in this analysis. Although 37 listeners indicated on the questionnaire that they were interested in uploading a clip of themselves reading "Comma gets a cure," only 12 listeners did so.

multiple predictors on regional identification plus 12 scales, which would mean well over 100 significance tests if this study used frequentist methods. In short, for this study to use null-hypothesis significance testing would be to invite false positives.[35] Bayesian inference, by contrast, does not rely on the notion underlying *p* value calculation of counterfactual samples that may unduly elicit false positives. The distribution of credible parameter values that results from a Bayesian analysis does not change based on the number of inferences that are drawn from it, nor do additional inferences change the probability of false positives (Kruschke 2015).

Bayesian inference also differs from frequentist inference in its end product; whereas frequentist methods produce a probability of observed data given a null hypothesis about the value of a given parameter (a *p* value), Bayesian methods produce a *posterior distribution*, which assigns different levels of credibility to possible values of given parameters. The posterior distribution is computed from the product of the *prior distribution*, which describes the researcher's pre-existing beliefs about distributions of possible values of the parameters, and the *likelihood function*, which describes the probability of observed data at different values of the parameters, divided by the sum of this product across all values of the parameters—or, in the case of continuously valued parameters, the integral of this product across all values of the parameters. While it is relatively straightforward to compute the denominator of the posterior

[35] The typical frequentist remedy for the inflated false-positive rate introduced by multiple comparisons is to apply a corrected significance level to individual significance tests, such as a Bonferroni correction, in which the overall significance level is divided by the number of intended tests. As Kruschke (2015) notes, however, two different experimenters may approach the same set of data with different numbers of intended significance tests, meaning they would apply different corrected significance levels to individual tests, possibly resulting in differing statements about the significance of some results. In other words, the corrected significance level approach is problematic because it means that conclusions about significance are partially a function of the experimenter's testing intentions. Moreover, in the present study's analysis, the application of a Bonferroni correction would mean a very low significance level—less than .0005—which would almost certainly invite false *negatives*.

distribution formula for simpler applications such as single discretely valued parameters,[36] the integral in the denominator is impossible to solve analytically for likelihood functions more complicated than the normal distribution; it is also possible to use numerical approximation methods to approximate the integral, but this strategy becomes computationally unwieldy when the model includes more than a handful of parameters (Kruschke 2015).

Instead, the most viable option for more complex models is to approximate the posterior distribution via Markov chain Monte Carlo (MCMC) probabilistic sampling. This method generates a sample of posterior values given the prior and likelihood by walking a 'chain' through possible values for the posterior. Over time, the chain will make the most visits to the most credible posterior values, meaning that the distribution of values visited by the chain will have its mode at the most credible posterior value. With a sufficiently large sample size—conventionally defined as an effective sample size (ESS) of 10,000—the distribution of values visited by an MCMC chain approximates the posterior distribution with tolerably low error (Kruschke 2015).

The data in this study were analyzed via Bayesian hierarchical models—models of data in which the parameters used to compute the lowest-level likelihood function are given prior distributions whose parameters (hyperparameters) are not fixed but are themselves estimated via higher-level prior distributions (Kruschke 2015).[37] These models were run in the statistical computing environment R (R Core Team 2015) by using the runjags package (Denwood 2015) to interface with the Bayesian sampling program JAGS (Plummer 2003); R scripts for running,

[36] Bayesian modeling can take multiple parameters into account simultaneously, and these parameters can have different prior distributions. The resulting posterior distribution allocates credibility to different values across multidimensional parameter space, and marginal posterior distributions can be computed for individual parameters.
[37] I outline the details of each hierarchical model in the next chapter (4.1. Regional identification and 4.2. Scales).

interpreting, and visualizing the results of MCMC runs were based on those of Kruschke (2015). Each simulation generated three parallel chains that ran on separate processors for the sake of efficiency, and chains were combined into one for the analysis of the posterior distribution. For each simulation, diagnostic plots (trace plots, density plots, Gelman plots, and autocorrelation plots) for selected parameters and hyperparameters were generated and inspected to ensure chains' convergence. In order to ensure sample sizes large enough to satisfactorily approximate the posterior, the sampling parameters of each model (i.e., number of burn-in steps, sample size, thinning interval) were adjusted so that the chains that sampled the parameters corresponding to each first-order predictor had an average ESS of at least 10,000.

Another common criticism of frequentist methods is that the use of *p* values encourages an outsize focus on binary significant/not significant decisions, rather than an accurate estimation of the magnitudes of effects (see e.g., Gelman et al. 2013:95). However, I find it useful to make inferential statements about predictors' effects on response variables by assessing whether the magnitude of a contrast between parameters corresponding to different levels of a predictor credibly falls above zero (akin to testing significance against a null hypothesis of zero). Rather than using a frequentist *p* value, this study uses a decision procedure based on the computation of 95% highest density intervals (HDIs), which indicate the "points of a distribution [that] are most credible, and which cover most of the distribution…such that every point inside the interval has higher credibility than any point outside the interval" (Kruschke 2015:87). This procedure works as follows: the difference between two parameter values, corresponding to different levels of a predictor (level A and level B), is computed at each step in the MCMC chain, and the distribution of these differences is a marginal posterior distribution (a *posterior contrast*) from which the 95% HDI is computed. If the 95% HDI of a posterior contrast between level A and

level B of the predictor excludes zero—that is, the HDI's lower limit is greater than zero or its upper limit is less than zero—then zero is not a credible value for the difference between responses in level A vs. level B; the difference between level A and level B is thus deemed to be a credible difference and the effect of the predictor is deemed to be a credible effect. Conversely, if the 95% HDI of the posterior contrast includes zero, then zero *is* a credible value for the difference between responses; the difference is thus not deemed to be credible—though this is not tantamount to saying that a lack of difference is credible (Kruschke 2015).[38,39] As a result, this study will make statements about the *credibility*, rather than significance, of certain effects and differences.

In the analysis of both regional identification and scales data, the main predictor of interest was guise: did listeners' responses change based on whether they heard a speaker with California-shifted vowels or non-shifted vowels? Other primary predictors were speaker-dependent factors: region, gender, and ethnicity. Posterior contrasts by guise are reported as conservative minus Californian; posterior contrasts by gender are reported as female minus male; posterior contrasts by ethnicity are reported as Latina/o minus Caucasian. Since region is a three-level predictor (Bay Area, Lower Central Valley, Southern California), region is reported as three pairwise posterior contrasts (fn. 38): Bay Area minus Lower Central Valley, Bay Area minus Southern California, and Lower Central Valley minus Southern California.

[38] One limitation of this decision procedure is that it can only indicate credible differences between pairs of levels, not three or more levels. As a result, statements about credible effects of factors with more than two levels (e.g., speaker region) must be based on multiple pairwise contrasts between levels. In the present study's analysis, multi-level factors are deemed to have a credible effect on predicted variables for which at least one pairwise contrast between levels is credibly nonzero.

[39] In contrast to Kruschke (2015), this study does not use a region of practical equivalence (ROPE) to make inferential claims about certain effects or differences. Since the scales data are reported in units of standard deviations (4.2. Scales), their effects are small in magnitude and not easy to interpret in terms of real differences, which makes the determination of a non-arbitrary ROPE practically difficult for scales data. The analysis of regional identification data also does not use a ROPE for consistency with the analysis of scales data.

Listener region was also analyzed as a predictor by restricting the analysis of listener region to the 83 listeners with hometowns in the Bay Area, Lower Central Valley, and Southern California, and their 497 trials. Two predictors were included only in the analysis of scales data: listeners' regional identification responses (i.e., did listeners' affective responses change based on where speakers were perceived to be from, regardless of where they were actually from?) and individual listeners. The analysis of scales data also considers interaction effects of various predictors (which was not possible for the regional identification analysis because of limitations of the model). Finally, 10 non-controlled predictors (beyond listener region, perceived region, and listener) were analyzed for unexpected effects: listener gender, listener ethnicity, listener mobility, listener age, listener mono/bi/multilingualism, listener experience with linguistics, student status, listening device, trial order, and number of listens.[40]

[40] Several non-controlled predictors contained factor levels with few listeners (e.g., 65 listeners self-identified as female, 31 as male, and one as genderqueer). As a result, the analysis of several non-controlled predictors excluded some trials from analysis and/or with collapsed factor levels. (The "excluded" trials were not actually excluded from the hierarchical models themselves but from the analysis of posterior contrasts; including these trials did not affect the credibility estimates for the effects that were included in the analysis.) The analysis of listener gender focused on 574 trials, including females and males, excluding a self-identified genderqueer listener. The analysis of listener ethnicity focused on 562 trials, including listeners who identified as Latina/o, Caucasian, Asian American, or multiple ethnicities, excluding African Americans and a listener who declined to state an ethnicity. The analysis of listener mobility focused on 562 trials, including listeners who had lived in the same city, same region of California, multiple regions of California, multiple US states (a combination of two categories: multiple Western states and multiple US regions), and multiple countries, excluding listeners of undetermined mobility. The analysis of listener experience with linguistics did not exclude any trials but collapsed levels of experience into two levels: some vs. no linguistic experience. The analyses of listener age (which was treated as a four-level categorical predictor), trial order (which was treated as a six-level categorical predictor), number of listens (which was treated as a five-level categorical predictor), listener mono/bi/multilingualism, student status, and listening device did not exclude any trials or collapse levels.

Chapter 4
Analysis

In this chapter, I describe the results of the analysis of the two main types of data yielded by the main study survey: regional identification data (map question) and language attitudes data (affective scales). I first discuss the regional identification responses (4.1. Regional identification), then the affective scales responses (4.2. Scales).

4.1. Regional identification

Across all 580 trials, speakers were identified as from the Bay Area in 24.5% of trials, the Lower Central Valley in 25.3% of trials, Southern California in 32.8% of trials, and outside California in 17.4% of trials. Despite no speakers being from outside California (making 17.4% of identification responses automatically incorrect), listeners correctly identified speakers' region in 172 trials, or 29.7% of all trials; this rate is consistent with Clopper and Pisoni's (2004) 30% accurate-recognition rate among six US regions and the pilot study's 25% accurate-recognition rate among four regions of California and one non-Californian region. This 29.7% rate does not accurately represent listeners' ability to differentiate Californian regions, however, as it includes

101 trials in which speakers were erroneously identified as non-Californians. Considering only the 479 trials for which listeners identified speakers as being from the Bay Area, Lower Central Valley, or Southern California, the accurate-recognition rate rises to 35.9%. This rate is not credibly greater than chance, as the posterior distribution of accurate-recognition rates, with a mode of 35.9% and 95% highest density interval (HDI) limits at 31.7% and 40.2%, does not exclude 33.3% as a credible value.[1] In other words, to the extent that Californian listeners accurately recognized Californian speakers' region, there is no evidence that this accurate recognition was anything more than luck.[2]

As the goal of this project is to describe the shape of folk belief rather than assess its accuracy, the following sections focus on how various predictors impacted the distributions of regional identification responses, not accurate-recognition rates. The analysis of these responses was carried out via Bayesian hierarchical models (see 3.4.6. Data analysis, above) in which distributions of responses in each category (e.g., each guise, each speaker gender) were modeled with a multinominal distribution as the likelihood function and Dirichlet distribution priors (see, e.g., Gelman et al. 2013:69). The priors for the individual proportions in the Dirichlet distributions were beta distributions with shape parameters 3 and 7 (representing a noncommittal prior assumption that the identification rate for each cell is 25%). The same one-predictor

[1] This posterior distribution is a beta distribution with shape parameters 175.64 and 313.36, which is directly calculated via Bayes' rule from a beta prior distribution with shape parameters 3.64 and 6.36 (representing a weak prior assumption that the accurate-recognition rate is 33.3%) and a Bernoulli likelihood function with 172 successes in a sample size of 479. (This simple application of estimating the posterior distribution of a single proportion can be directly calculated via Bayes' rule, rather than approximating the posterior via Markov chain Monte Carlo methods, the latter of which are used in the hierarchical models that perform the lion's share of the analysis in this chapter.)

[2] Paradoxically, the 29.6% accurate-recognition figure that includes trials in which speakers were identified as being non-Californians *is* credibly greater than chance (as directly calculated via the same Bernoulli-beta model described in fn. 1, except with a prior whose mode is 25% rather than 33.3%). In my opinion, however, this figure is not as meaningful as the 35.9% figure, since the lower figure includes outside-California responses that could not have been correct; the higher (but not credibly above chance) figure is a more accurate gauge of Californian listeners' ability to differentiate these regions.

hierarchical model structure was used for all predictors individually: the primary predictors (guise, speaker region, speaker gender, and speaker ethnicity), listener region, and other non-controlled predictors. Because of limitations of this model, interaction effects could not be analyzed.

The following sections describe regional identification responses as affected by several predictors: speakers' use of the California Vowel Shift (4.1.1. Guise), other primary predictors (4.1.2. Speaker region, gender, and ethnicity), listeners' regional origin (4.1.3. Listener region), and other non-controlled predictors such as listener mobility and trial order (4.1.4. Non-controlled predictors). I also briefly discuss listeners' responses when asked to specify a regional identification response of "outside California" (4.1.5. Write-in blanks). I conclude by summarizing the regional identification results (4.1.6. Summary: Regional identification).

4.1.1. Guise

Figure 4.1 displays the distributions of regional identification for the conservative and Californian guises, which differed only by speakers' use or non-use of the California Vowel Shift (CVS). Figure 4.1 suggests that there was basically no difference between the conservative and Californian guises in the rate at which speakers were identified as being from the Bay Area or Lower Central Valley, but that more California-shifted speakers were identified as Southern Californian and more non-shifted speakers were identified as non-Californian. The conservative minus Californian posterior contrasts for each response category, which are displayed in Figure 4.2 along with their modes and 95% highest density intervals (HDIs), support these observations, as for Southern California the 95% HDI is completely below zero and for outside California the 95% HDI is completely above zero.

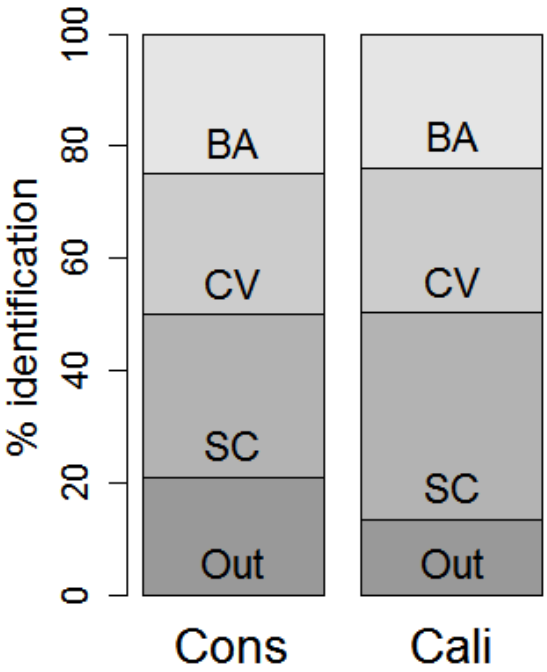


Figure 4.1. Regional identification by guise: Cons[ervative] vs. Cali[fornian].

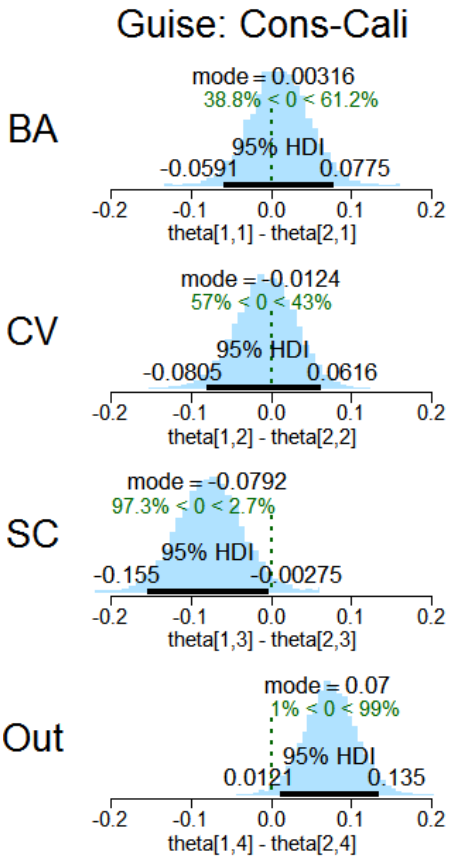


Figure 4.2. Posterior regional identification contrasts by guise: Cons[servative] minus Cali[fornian].

4.1.2. Speaker region, gender, and ethnicity

Figure 4.3 displays the distributions of regional identification by speaker region, with pairwise posterior contrasts in Figure 4.4. As Figure 4.4 indicates, speaker region had a credible effect on the rates at which speakers were identified as being from Southern California or outside California, but not the Bay Area or Lower Central Valley. Speakers from the Lower Central Valley and Southern California were identified as being from Southern California at a credibly higher rate than speakers from the Bay Area, and speakers from the Bay Area were identified as being from outside California at a credibly higher rate than speakers from the Lower Central Valley.

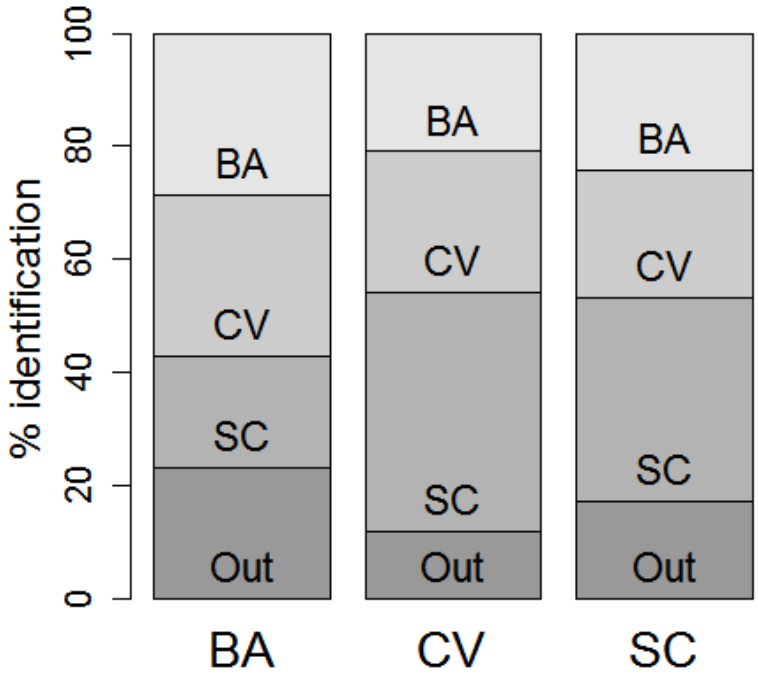


Figure 4.3. Regional identification by speaker region.

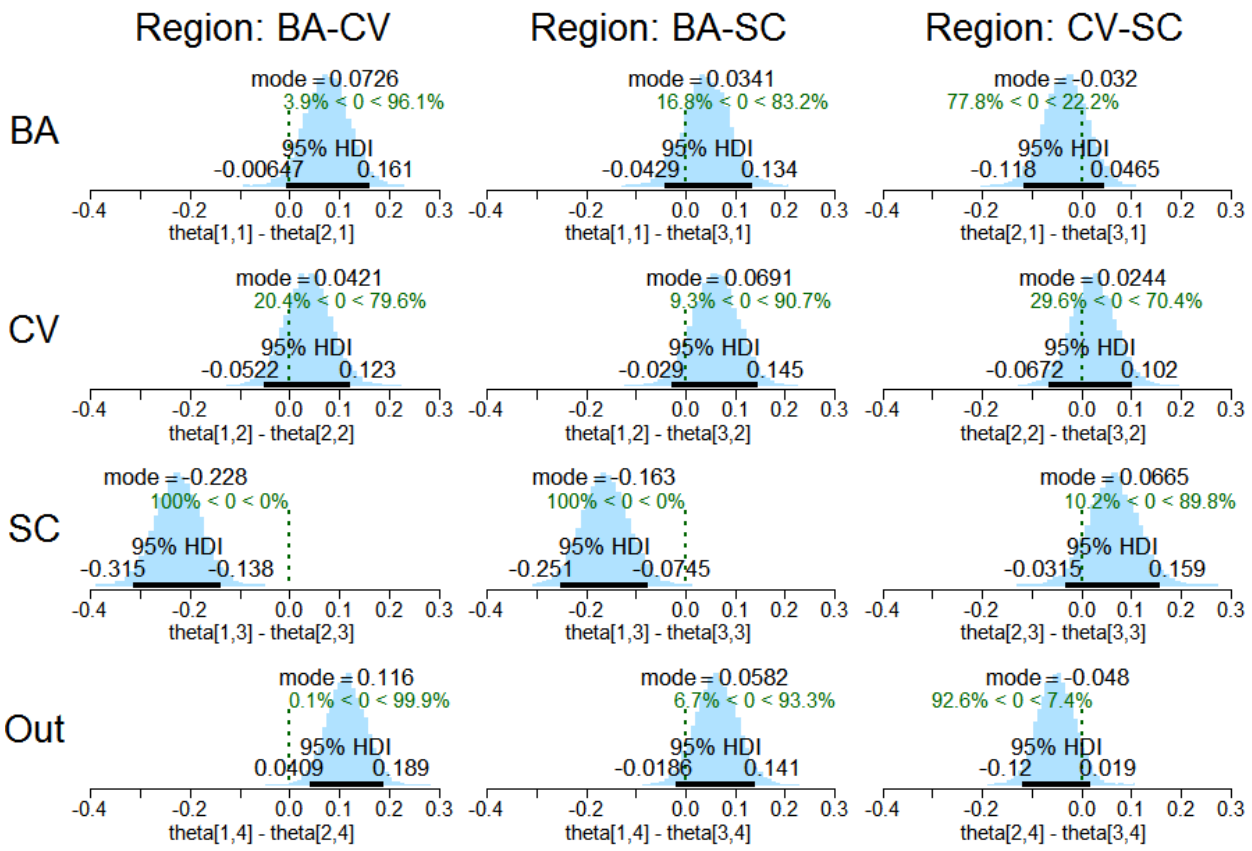


Figure 4.4. Pairwise posterior regional identification contrasts by speaker region.
The pairwise contrasts are Bay Area minus Lower Central Valley, Bay Area minus Southern California, Lower Central Valley minus Southern California.

The distributions of regional identification by speaker gender displayed in Figure 4.5 suggest that male speakers were more likely to be identified as being from the Bay Area and female speakers were more likely to be identified as being from Southern California. These observations are supported as credible differences by the posterior contrasts in Figure 4.6.

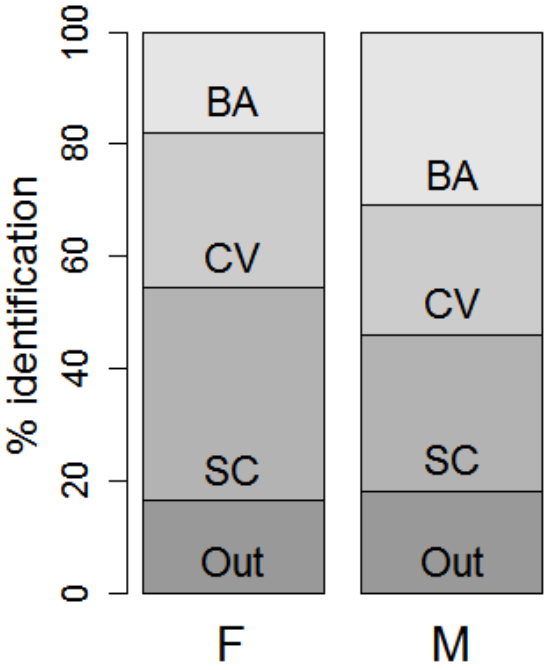


Figure 4.5. Regional identification by speaker gender: F[emale] vs. M[ale].

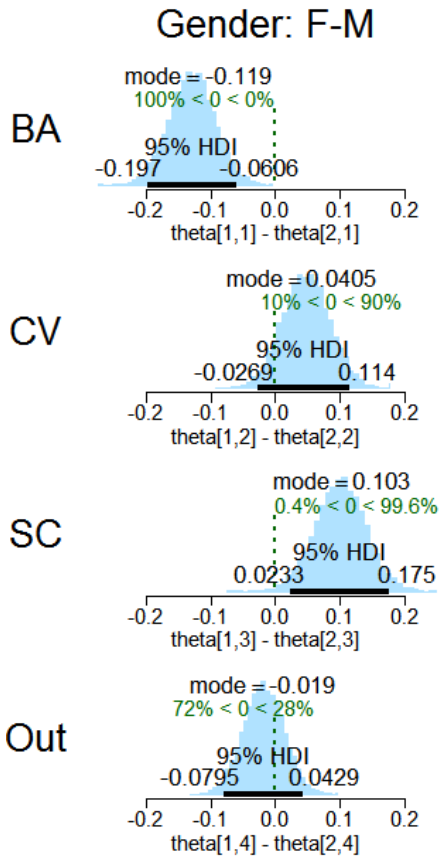


Figure 4.6. Posterior regional identification contrasts by speaker gender: F[emale] minus M[ale].

As Figure 4.7 indicates, the distributions of regional identification were nearly identical for Latina/o and Caucasian speakers. The posterior contrasts by ethnicity in Figure 4.8 indicate no credible effect of ethnicity on regional identification.

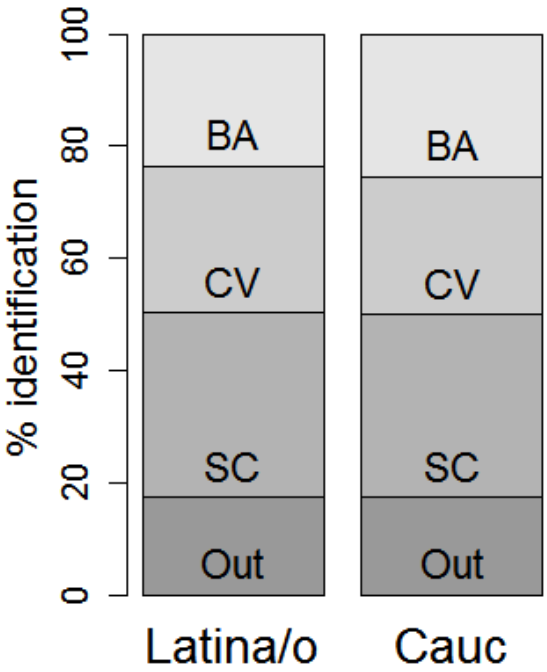


Figure 4.7. Regional identification by speaker ethnicity: Latina/o vs. Cauc[asian].

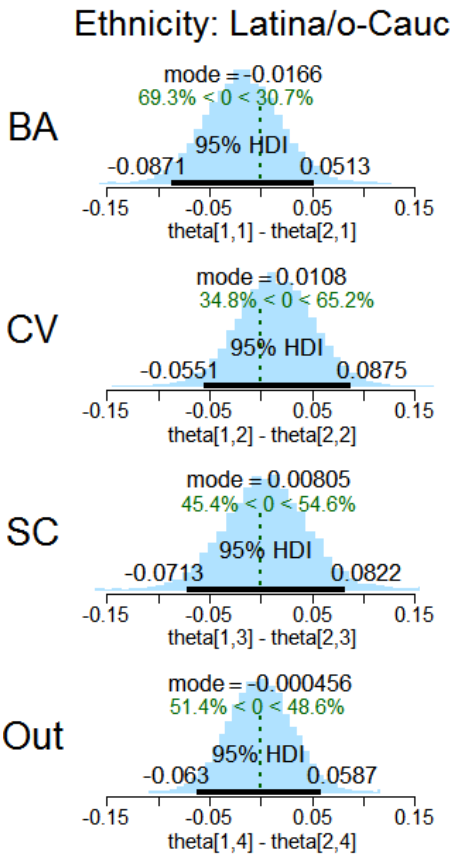


Figure 4.8. Posterior regional identification contrasts by speaker ethnicity: Latina/o minus Cauc[asian].

To summarize, regional identification was credibly affected by guise, speaker region, and speaker gender, with no credible effect of ethnicity. Speakers were more likely to be identified as being from Southern California if they used CVS features, were female, or were from the Lower Central Valley or Southern California. Speakers were more likely to be identified as being from outside California if they did not use CVS features or were from the Bay Area. Speakers were more likely to be identified as being from the Bay Area if they were male. Notably, none of the primary predictors influenced the rate at which speakers were identified as being from the Lower Central Valley. It is possible that this result further indicates the erasure of Lower Central Valley speakers that was found in the pilot (see 2.3. Discussion, above).

4.1.3. Listener region

Figure 4.9 displays the distributions of regional identification by listener region. The patterns in Figure 4.9 suggest local preferences among Lower Central Valley and Southern California listeners, who identified speakers as from their home region more than anywhere else; among Bay Area listeners, the Bay Area and Southern California tied for highest identification rate, with 31.0% of trials apiece. The pairwise posterior contrasts displayed in Figure 4.10 indicate that these local preferences are indeed credible differences; listeners from each region were credibly more likely to identify speakers as from their home region than listeners from either other region (with a weaker effect for Southern California, for which only the difference in rates of identification between Lower Central Valley and Southern California listeners was credibly nonzero). Bay Area listeners were also credibly more likely than Lower Central Valley listeners to identify speakers as being from outside California.

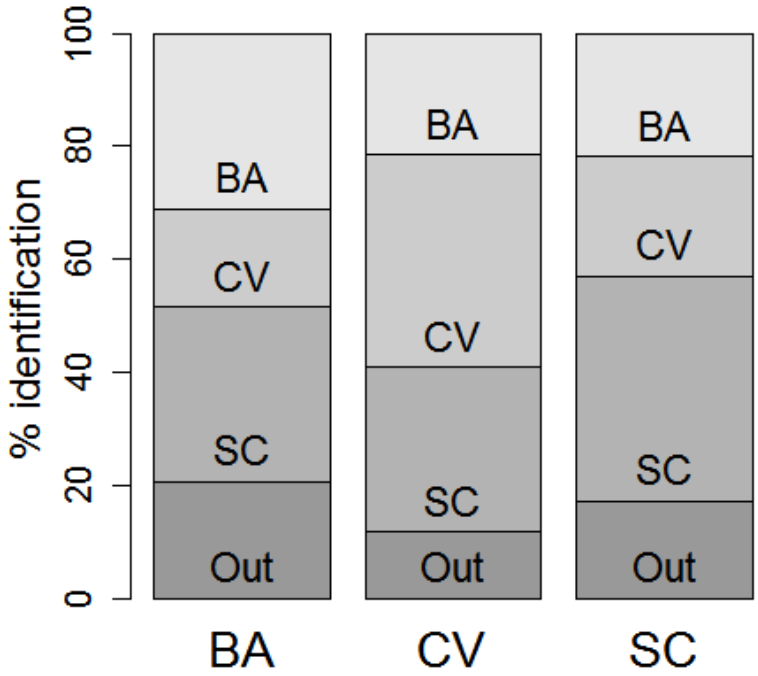


Figure 4.9. Regional identification by listener region.

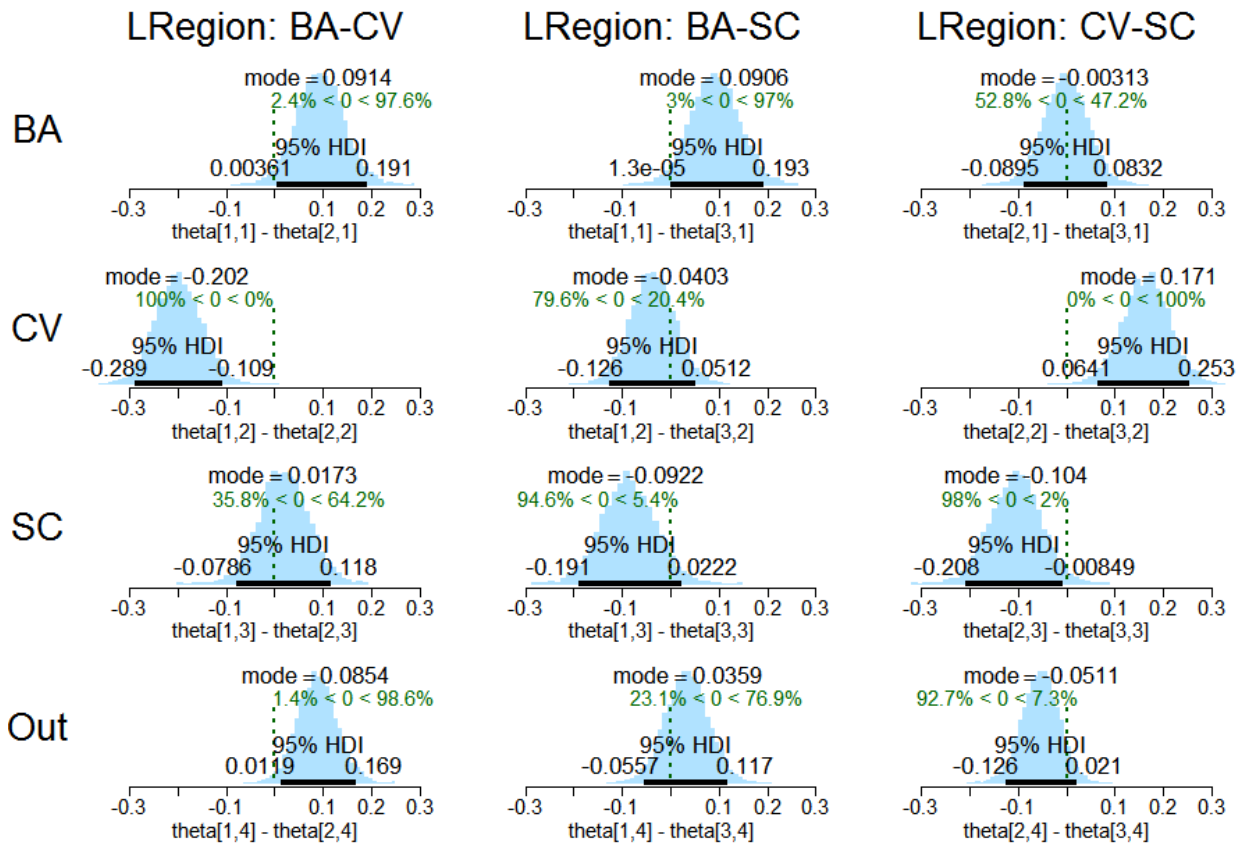


Figure 4.10. Pairwise posterior regional identification contrasts by listener region.

It is also worth considering how listeners from different regions responded differently to the conservative and Californian guises by looking at the distribution of conservative minus Californian contrasts in regional identification responses by listener region. For example, Bay Area listeners identified speakers as being from Southern California at a rate of 27.5% in the conservative guise and 34.7% in the Californian guise, a change of 7.2 percentage points. These contrasts are displayed in Table 4.1.[3]

	Regional identification			
Listener region	**BA**	**CV**	**SC**	**Out**
BA	−2.0	2.8	−7.2	6.4
CV	−0.8	−1.9	−2.4	5.1
SC	7.5	−1.4	−12.8	6.8

Table 4.1. Percentage-point change in regional identification by listener region.
Changes are represented as conservative minus Californian. Positive values indicate greater percentage of conservative; negative values indicate greater percentage of Californian.

Several interesting trends emerge from these contrasts. First, listeners from each region identified the Californian guise as being from outside California at a lower rate than the conservative guise; that is, the overall pattern displayed in Figure 4.1 of identifying the Californian guise as from outside California at a credibly lower rate is found consistently across regions. Likewise, listeners from each region gave the highest increase in identification rates to Southern California, which is also in keeping with the overall pattern in Figure 4.1. While Bay Area and Lower Central Valley listeners identified the Californian guise as members of their regional ingroup at a higher rate than the conservative guise, this increase was greater for speakers whom these listeners perceived to be from Southern California. In short, listeners across

[3] Note that these are merely descriptive statistics that do not permit statements about credible effects. The trends that they indicate are nevertheless compelling.

regions of California agreed that California-shifted speakers sounded more like they were from Southern California and less like they were from outside California.

In summary, the regional identification results by listener region indicate a preference among listeners from across California for identifying speakers across the board as from their respective home regions. These results also reinforce the pattern introduced in the analysis of regional identification by guise (4.1.1. Guise) that suggests Southern California is especially associated with the CVS—perhaps especially by Southern California listeners themselves.

4.1.4. Non-controlled predictors

Four non-controlled predictors had no credible effect on regional identification: listener age, listener mono/bi/multilingualism, listener experience with linguistics, and listening device. Six non-controlled predictors yielded at least one credible difference on regional identification: trial order (five credible differences), listener mobility (five credible differences), listener ethnicity (four), number of listens (three), listener gender (one), and student status (one). Most of these differences involved the rates at which speakers were identified as being from the Lower Central Valley, and the listener demographics that favored identifying speakers as from the Lower Central Valley were also those for which Lower Central Valley listeners were more heavily represented in the listener sample. Aside from these differences, however, non-controlled predictors did not substantially affect regional identification overall.

Trial order was analyzed to assess whether listeners' regional identification responses changed over the course of the listening task in a systematic way. Trial order yielded five credible differences (with six trials, there were 15 pairwise contrasts for each of the four regions), all of which involved trial 5. Speakers were identified as from the Lower Central

Valley at a credibly lower rate in trial 5 than in trial 2, trial 4, and trial 6, and as from Southern California at a credibly higher rate in trial 5 than in trial 3 and trial 4. It is not clear why speakers in trial 5 were associated with Southern California and not the Lower Central Valley (especially given that no primary predictors credibly influenced Lower Central Valley identification). Otherwise, however, the distributions of regional identification did not systematically change over the course of the listening task.

Listener mobility yielded five credible differences (with 10 pairwise contrasts for each of the four perceived regions). These differences suggested that less mobile listeners were more likely to identify speakers as being from the Lower Central Valley and less likely to identify speakers as being from outside California. In particular, speakers were more likely to be identified as from the Lower Central Valley by listeners who had lived in the same city than listeners who had lived in multiple regions of California, multiple US states, or multiple countries; listeners who had lived in multiple cities within the same Californian region were more likely than those who had lived in multiple US states to identify speakers as from the Lower Central Valley; and listeners who had lived in the same city were less likely than those who had lived in multiple US states to identify speakers as from outside California. As discussed below, this effect likely reflects the preponderance of Lower Central Valley listeners in the lower categories of mobility, rather than a larger trend of less-mobile listeners being more likely to hear speakers as being from the Lower Central Valley.

Listeners self-reported the number of times they listened to each clip for each trial. The distribution of self-reported listens skewed high, with a median of 2 listens and a mean of 2.5 listens per trial; the counts of trials with more than five listens were in the single digits for each number of listens. Number of listens was analyzed as a categorical factor with five levels: one

listen (28.1% of trials), two listens (34.1% of trials), three (20.2%), four (9.5%), and five or more listens (8.1%); with five levels, the analysis of number of listens involved 10 pairwise contrasts for each of the four regions. Number of listens yielded three credible differences, all involving the four-listens category. Trials in which the listener heard the clip four times were credibly more likely to result in an identification of outside California than one listen or two listens, and credibly less likely to result in an identification of the Lower Central Valley than three listens. In other words, the more that listeners heard a clip, the less likely they were to identify the speaker as being from the Lower Central Valley and the more likely they were to identify the speaker as being from outside California.[4]

Listener ethnicity yielded four credible differences (with six pairwise contrasts for each of the four regions), all involving Latina/o listeners and Caucasian or Asian American listeners. Latina/o listeners were credibly more likely than Caucasian listeners and Asian American listeners to identify speakers as from the Lower Central Valley. Latina/o listeners were also less likely than Caucasian listeners and Asian American listeners to identify speakers as from Southern California. Listener gender and listener student status each yielded just one credible difference (with two levels apiece, both predictors involved one pairwise contrast for each of the four regions). With respect to listener gender, female listeners identified speakers as from outside California at a greater rate than male listeners. With respect to listener student status, listeners who were students were credibly more likely than non-students to identify speakers as from the Lower Central Valley.

[4] It is unlikely that this effect is an indirect effect of listener region; if so, then Lower Central Valley listeners should have had more listens than other regions, but Lower Central Valley listeners actually had a slightly lower mean number of listens than did Bay Area listeners (2.49 vs. 2.58).

In general, non-controlled predictors affected Lower Central Valley identification more than other perceived regions, with credibly higher rates of Lower Central Valley identification among Latinas/os, students, and listeners who had lived in the same city their entire lives. Each of these demographics were also represented more heavily among Lower Central Valley listeners than the other listener regions. Forty-two percent of Lower Central Valley listeners had lived in the same city, as compared with just 3.8% of Bay Area and Southern California listeners; Latinas/os were 38.7% of Lower Central Valley listeners, 19.2% of Southern California listeners, and 11.5% of Bay Area listeners; and students were 93.5% of Lower Central Valley listeners, 76.9% of Southern California listeners, and 46.2% of Bay Area listeners. In other words, the effects of non-controlled predictors may simply be a reflection of the previous section's finding that Lower Central Valley listeners were more likely to identify speakers across the board as being from the Lower Central Valley (4.1.3. Listener region). This result is also interesting in light of the fact that none of the primary predictors credibly affected the rate at which speakers were identified as being from the Lower Central Valley.

4.1.5. Write-in blanks

In all 101 of the trials in which listeners identified speakers as being from outside California, listeners filled in the blank specifying where they thought the speaker was from, although in 25 trials listeners indicated that they were unsure of the speaker's regional origin. The responses to the other 76 trials were coded for regions of the US or outside the US: Northeast (30), Midwest (13), South (8), Other US (6), Other West (5), and International (14). Guise and speaker ethnicity credibly affected the rate at which these speakers were identified as International: 4 of 63 conservative speakers (6.3%) vs. 10 of 38 Californian speakers (26.3%), and 13 of 48 Latina/o speakers (27.1%) vs. 1 of 53 Caucasian speakers (1.9%) were identified as

International. In addition, 21 of 45 Bay Area speakers were identified as from the Northeast (46.7%), as compared with 2 of 24 Lower Central Valley speakers (8.3%) and 7 of 32 Southern California speakers (21.9%); pairwise contrasts revealed credible differences between the Bay Area and Lower Central Valley and between the Bay Area and Southern California.[5] Speaker gender did not credibly affect responses for any category.

In sum, when listeners perceived speakers as being from outside California, their perception of exactly where speakers were from was shaped by guise, speaker region, and speaker ethnicity. Californian-shifted speakers and Latinas/os were more likely to be perceived as from outside the US, and Bay Area speakers were more likely to be perceived as from the Northeast.

4.1.6. Summary: Regional identification

The analysis of regional identification responses revealed several noteworthy results. Among trials in which listeners identified speakers as being from one of the three focal regions of California, listeners accurately recognized speakers' region in 35.9% of trials, which is not credibly greater than chance. Regional identification was sensitive to guise, as listeners associated the CVS with Southern California and non-CVS features with non-Californians. Speaker gender also influenced regional identification, with females more likely to be identified as Southern Californians and males more likely to be identified as Bay Area speakers; speaker ethnicity, by contrast, had virtually zero effect on regional identification. The analysis of listener region indicated local preferences among each listener region, though listeners in each region agreed that the CVS was especially associated with speakers believed to be from Southern

[5] These differences were modeled via an MCMC simulation with a Bernoulli likelihood function and a noncommittal beta prior with shape parameters 0.5 and 0.5. The ESSs of the parameters modeling the proportions averaged at least 10,000.

California. None of the primary predictors credibly affected the rate at which speakers were identified as from the Lower Central Valley (although a local preference among Lower Central Valley listeners was reflected in several credible effects of non-controlled predictors) possibly indicating that the Lower Central Valley is not highly salient to Californian listeners—even with a listener sample that included Lower Central Valley speakers. Finally, the analysis of listeners' responses to the outside California write-in blanks suggested that when Californians hear CVS users and/or Latina/o speakers as non-Californians, they are more likely to hear them not merely as non-Californians but foreigners.

4.2. Scales

Scales data were first standardized to control for listeners' differential use of the slider bars; the average listener used a range of 92.8 (out of a maximum 100), and some listeners used far less. Standardized scores were computed by subtracting individual listeners' mean scores from their raw scores and dividing by individual listeners' standard deviations. All results are reported here in unit-less standard deviations (which means in many cases graphs do not exceed −1 or +1 on the response axis). Across listeners, the overall mean score was 55.3 and the overall standard deviation was 25.1; a difference of 1 standard deviation can thus be interpreted as a difference of roughly one quarter of the slider bar for the average listener in this survey.

Principal component analyses (PCAs) were performed on scales data to assess the viability of reducing the 12 scales into a smaller set of response variables, as well as to expose relationships between scales. A PCA on the entire data set failed to yield a set of principal components that accounted for a satisfactorily large proportion of the variance. The data set was then divided into subsets by speaker gender since *feminine* was unlikely to pattern with other scales in the same way for both female and male speakers. PCAs on these subsets also failed to

yield principal components that accounted for a satisfactorily large proportion of the variance, however. As a result, all 12 scales were analyzed separately as response variables (i.e., each scale was modeled separately).

The analysis of scales data was carried out via two different Bayesian hierarchical models for different sets of predictors. The primary analysis was carried out via an eight-predictor model that included the primary predictors (guise, speaker region, speaker gender, and speaker ethnicity), listener region, perceived region (i.e., listeners' regional identification of speakers), and listener and speaker (which were not analyzed but included as a 'random effect'[6]). This model included nine second-order predictors to assess interaction effects: guise × speaker region, guise × speaker gender, guise × speaker ethnicity, guise × listener region, guise × perceived region, speaker region × speaker gender, speaker region × speaker ethnicity, speaker gender × speaker ethnicity, and listener region × perceived region. Speaker was included as a third-order interaction of speaker region × speaker gender × speaker ethnicity, as this interaction is equivalent to speaker in this data.[7] Second, the effects of non-controlled predictors were analyzed via individual two-predictor models that included speaker, but not listener, as an un-analyzed 'random effect.'[8] For all models, the chains that sampled the parameters corresponding

[6] Bayesians tend to enclose 'random effect' in scare quotes; unlike frequentist linear models, which treat some effects as fixed effects and others as random effects, Bayesian analysis conceptualizes *all* effects as random effects (Gelman & Hill 2006:3). The motivation behind including participant (in this case, listener) as a predictor in Bayesian analysis is the same motivation behind including participant as a random effect in linear modeling: making inferences beyond the participants who were sampled.

[7] Whereas frequentist models of categorical predictors with numerical response data (i.e., ANOVA) require the inclusion of all interaction terms in the model (e.g., second, third, fourth, and fifth-order interactions in a five-predictor model), Bayesian modeling does not obligate the experimenter to include any interaction terms beyond that which the experimenter seeks to analyze (Kruschke 2015). Moreover, I found that including interactions beyond the third order caused MCMC chains to be badly autocorrelated, making it impossible to achieve ESSs even approaching 10,000. Since I am not interested in interactions beyond the second order, the multi-predictor hierarchical model only included second-order interactions (save for the third-order effect standing in for speaker).

[8] Listener could not be included as a 'random effect' in these two-predictor models of non-controlled predictors since almost all non-controlled predictors were listener-dependent (e.g., listener ethnicity, listening device). Multi-predictor hierarchical models do not need all interaction cells to be filled (Kruschke 2015), a fact that allowed for

to each first-order predictor (except for listener, in the eight-predictor model) had an average ESS of at least 10,000.

All of these models modeled scale ratings with a normal distribution as the likelihood function and various priors on the hyperparameters of this likelihood function.[9] Preliminary simulations used a Student's *t* distribution as the likelihood function, as normality of the responses was not assumed; the posterior distributions of the normality parameter were consistently high in these simulations, however, so a normal likelihood was used instead to increase efficiency. At each step in the chain, these models computed the baseline response level (i.e., the mean value of each combination of factor levels) as well as the deflection from the baseline contributed by each factor level, such that the deflections of levels within each factor summed to zero. The following results are reported in terms of these deflections from the baseline (and especially contrasts between deflections), in units of standard deviations, as a means of assessing the effects of different predictors.

In the following sections, I first describe the scales responses by primary predictors, guise (4.2.1. Guise) and others (4.2.2. Speaker region, gender, and ethnicity), and summarize these main effects (4.2.3. Summary: Primary predictors) before describing how primary predictors interacted with one another (4.2.4. Interactions of primary predictors). I then discuss the effects

---

the inclusion of both listener and listener region in the eight-predictor model. However, the distribution of empty cells in models including listener alongside a non-controlled predictor led to unreliable results; as a result, listener was excluded from these models.

[9] In particular, the mean and standard deviation of the normal likelihood function were also modeled as hyperparameters. Mean hyperparameters were modeled as the sum of a first-order intercept (which was modeled by a normal prior) and individual predictor terms for the relevant combination of factor levels. For the eight-predictor model, these predictor terms were first-order and second-order predictors, with a third-order predictor corresponding to speaker; for the two-predictor models of non-controlled predictors, these predictor terms were first-order predictors. Predictor terms were modeled individually by normal priors with a mean of zero. Standard deviations for each combination of factor levels were modeled by a gamma prior whose shape and rate parameters were modeled by gamma priors with a mode equal to half the (observed) standard deviation of the predicted variable and a standard deviation equal to twice the (observed) standard deviation of the predicted variable.

of listener region and perceived region as main effects (4.2.5. Listener region and perceived region), as they interacted with guise (4.2.6. Interactions of guise with listener region and perceived region), and as they interacted with one another (4.2.7. Interaction of listener region with perceived region). As with regional identification responses, I outline the effects of non-controlled predictors such as listening device (4.2.8. Non-controlled predictors). I conclude this chapter by summarizing the results of the analysis of scales (4.2.9. Summary: Scales).

4.2.1. Guise

Guise credibly affected responses to three scales, *Californian*, *confident*, and *sounds like a Valley girl*, as California-shifted speakers were rated higher than non-shifted speakers on all three scales. As Figure 4.11 indicates, all of these credible differences are small in magnitude; the largest credible difference for *Californian*, for example, is a 0.39 standard deviation increase for the Californian guise above the conservative guise, roughly corresponding to a difference of 10 points on the 0–100 slider bar.[10]

[10] Although Figure 4.11 appears to show the *confident* posterior contrast as having a 95% HDI with an upper bound at zero, this is a rounding error; the actual upper bound is -3.6×10^{-3}. Other apparent zeroes in other plots (e.g., Figure 4.15) are also rounding errors.

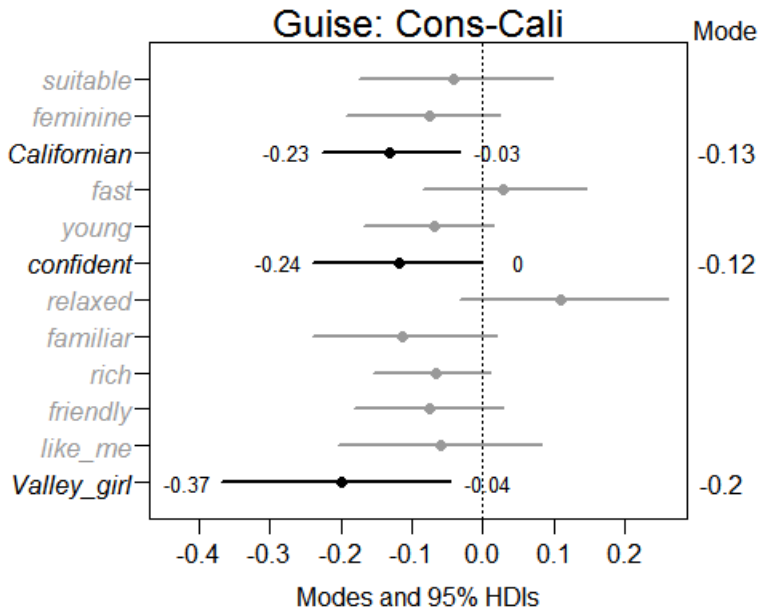


Figure 4.11. Modes and 95% HDIs of posterior scales contrasts by guise: Cons[servative] minus Cali[fornian].
Black labels and bars indicate scales with contrasts credibly above or below zero (i.e., credible effect of guise on that scale). Gray labels and bars indicate contrasts not credibly above or below zero (i.e., no credible effect of guise on that scale). This graph, and the ones like it in this chapter, are essentially condensed representations of the type of posterior contrast histogram plots in the Regional identification section.

The fact that both *Californian* and *sounds like a Valley girl* were rated higher for the Californian guise could lead us to believe that these are interrelated or equivalent perceptual constructs to Californian perceivers. However, as Figure 4.12 indicates, the higher ratings for the Californian guise are relative to different baseline ratings: a high baseline for *Californian* and a low baseline for *Valley girl*. In other words, although listeners generally heard speakers across all trials as Californian and not like Valley girls, the California-shifted speakers sounded even *more* Californian and the non-shifted speakers sounded even *less* like Valley girls.

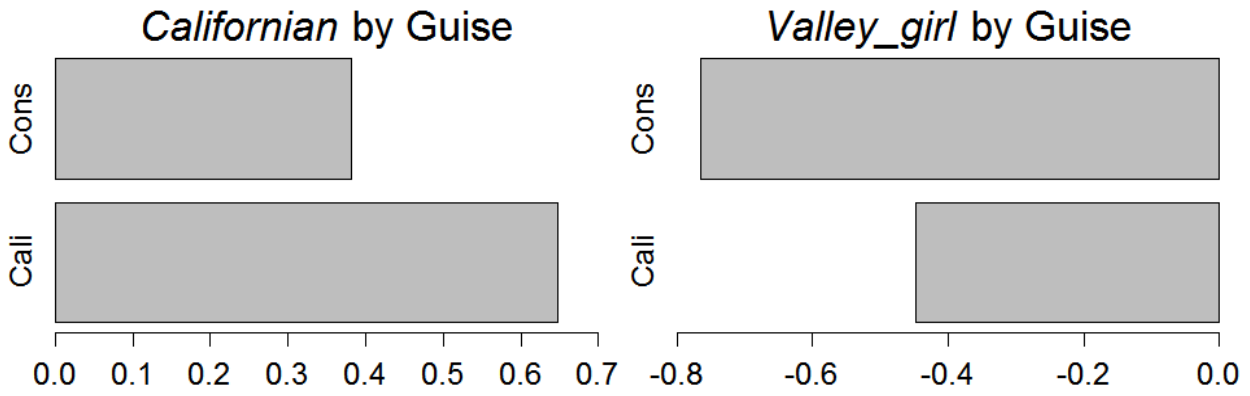


Figure 4.12. Mean standardized *Californian* and *sounds like a Valley girl* ratings by guise.

4.2.2. Speaker region, gender, and ethnicity

Speaker region had a credible effect on responses to eight scales: all but *Californian*, *relaxed*, *familiar*, and *sounds like me*. As Figure 4.13 indicates, Bay Area speakers were rated credibly higher than both Lower Central Valley and Southern California speakers on *suitable*, *confident*, and *rich*. Lower Central Valley speakers were rated highest of any region on *young* and *sounds like a Valley girl*. Southern California speakers were rated highest of any region on *friendly* and *feminine*, and lowest on *fast*.

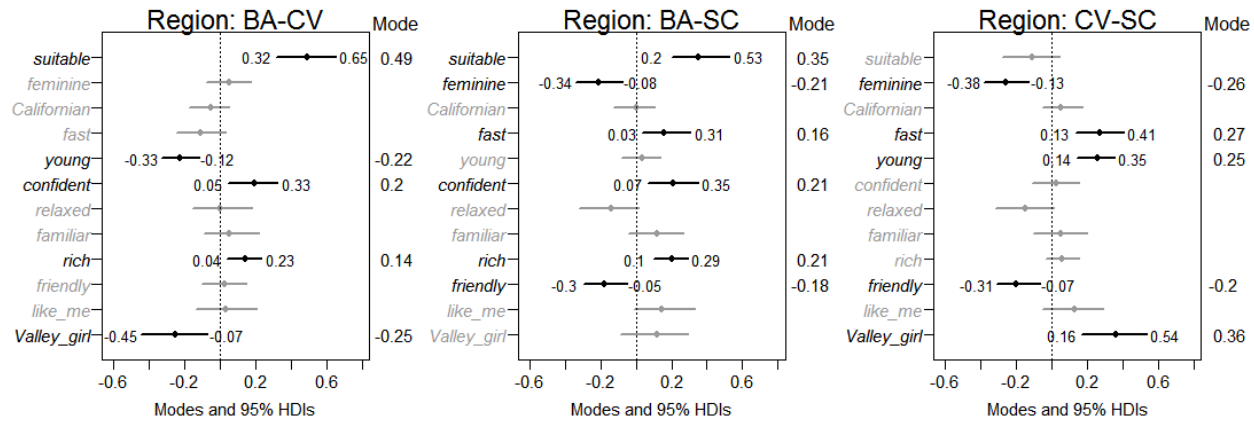


Figure 4.13. Modes and 95% HDIs of pairwise posterior scales contrasts by speaker region.
The contrasts are Bay Area minus Lower Central Valley, Bay Area minus Southern California, Lower Central Valley minus Southern California.

As Figure 4.14 indicates, speaker gender had a credible effect on responses to eight scales, with female speakers rated credibly higher on *feminine*, *Californian*, *relaxed*, and *sounds*

like a Valley girl, and male speakers rated credibly higher on *suitable*, *fast*, *confident*, and *friendly*. Speaker gender had a greater effect on *feminine* and *Valley girl* than any other primary predictor had on any other scales. It is not surprising that female speakers were rated substantially higher on these scales—the least credible difference for *feminine* is almost 2 standard deviations, or nearly half of the slider bar—as these scales are inherently gendered and speakers' gender was transparent to listeners in a way that guise, region, or ethnicity were not.

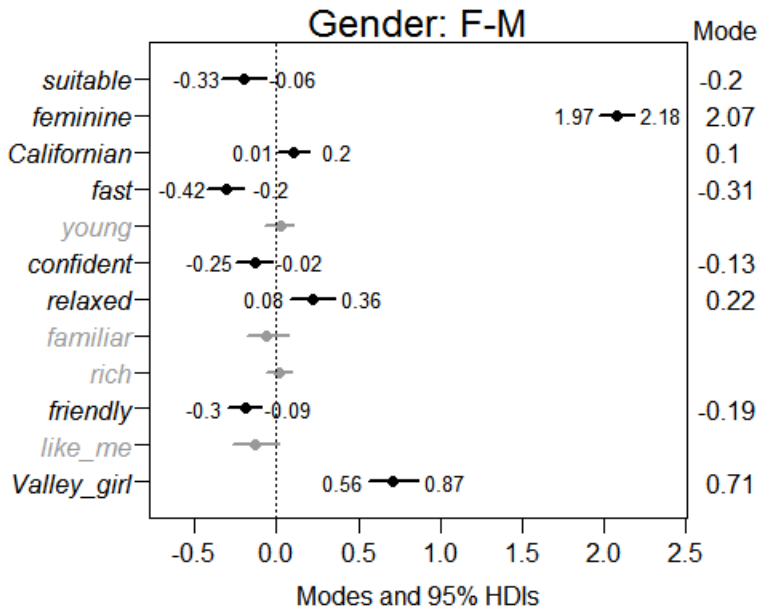


Figure 4.14. Modes and 95% HDIs of posterior scales contrasts by gender: F[emale] minus M[ale].

As Figure 4.15 indicates, speaker ethnicity had a credible effect on eleven scales, with Latina/o speakers rated credibly higher on *feminine* and *relaxed*, and Caucasian speakers rated credibly higher on *suitable*, *Californian*, *fast*, *young*, *confident*, *familiar*, *rich*, *friendly*, and *sounds like me*. Interestingly, *sounds like a Valley girl* was the only scale for which ethnicity did not credibly affect ratings, despite the fact that the Valley girl character arguably represents a Caucasian stereotype and thus comes closer to being inherently raced than does any other scale.

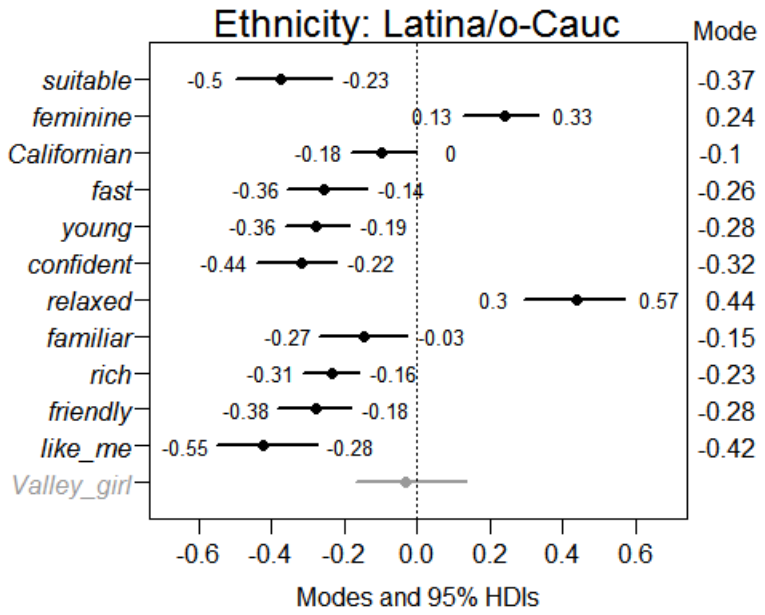


Figure 4.15. Modes and 95% HDIs of posterior scales contrasts by ethnicity: Latina/o minus Cauc[asian].

4.2.3. Summary: Primary predictors

Table 4.2 summarizes the results by guise, speaker region, speaker gender, and speaker ethnicity by indicating the credible effects of each primary predictor. It is clear from the bottom row of Table 4.2 that guise had a smaller effect than the other primary predictors on these scales, as it resulted in credible differences on just three scales. This small effect of guise is also clear from the posterior contrasts of the three credible differences by guise, as 0.05 is included as a credible value in their 95% HDIs; in other words, the effect of guise on *Californian*, *confident*, and *Valley girl* may be as small as 1.25 notches on the 0–100 slider bar. Although these scales are by no means representative of all of the responses that Californian listeners may have to language variation—that is, given a different set of scales, guise could conceivably yield a greater number of credible effects than other predictors—these results suggest that when Californian listeners react to speakers, the presence or absence of the CVS does not register as

strongly as other speaker characteristics. Speaker ethnicity, on the other hand, registers rather strongly in listeners' ratings.

Scale	Guise	Region	Gender	Ethnicity	Credible effects
suitable		BA > CV, SC	M > F	Cauc > Latina/o	**3**
feminine		SC > BA, CV	F > M	Latina/o > Cauc	**3**
Californian	Cali > Cons		F > M	Cauc > Latina/o	**3**
fast		BA, CV > SC	M > F	Cauc > Latina/o	**3**
young		CV > BA, SC		Cauc > Latina/o	**2**
confident	Cali > Cons	BA > CV, SC	M > F	Cauc > Latina/o	**4**
relaxed			F > M	Latina/o > Cauc	**2**
familiar				Cauc > Latina/o	**1**
rich		BA > CV, SC		Cauc > Latina/o	**2**
friendly		SC > BA, CV	M > F	Cauc > Latina/o	**3**
like me				Cauc > Latina/o	**1**
Valley girl	Cali > Cons	CV > BA, SC	F > M		**3**
Credible effects	**3**	**8**	**8**	**11**	

Table 4.2. Credible effects by primary predictor and scale.
X > Y indicates credibly higher rating for X than Y. Blank cells indicate no credible effect.

The results by speaker region indicate substantial differences from the pilot study, in which Southern California speakers received the most favorable rating on each scale. In the main study, Southern California speakers were rated credibly highest on *friendly* but credibly lowest on *fast*; in the pilot, by contrast, Southern California speakers received the highest ratings of any region on *fast*.[11] Similarly, Bay Area speakers were also rated differently in the main study and the pilot, with the pilot Bay Area speakers receiving the lowest ratings of any region on *confident* but the main study Bay Area speakers receiving the highest ratings of any region on *confident*. These results, along with listeners' relatively low accurate-recognition rate, suggest that speaker region is not a reliable basis for listeners' ratings—that is, the high ratings for Southern

[11] Southern California's bottom rating on *feminine* is also a top rating on *masculine*, and neither pole of this scale is easily interpretable as more favorable in the way that *confident* or *friendly* is.

California speakers in the pilot study and low ratings for Southern California speakers in the main study were not a property of any defining Southern California speech characteristics (or any change in Southern California's status in the intervening time between these two studies), but rather an effect of the specific speakers who represented Southern California in these studies. However, as the analysis of perceived region as a predictor shows, listeners' ratings were also impacted by the *idea* of what speakers from certain regions sound like (see 4.2.5. Listener region and perceived region, below).

It is also notable that the Bay Area and Southern California share the same side of just two credible differences, *young* and *sounds like a Valley girl*; for the other six credible effects of region, the Lower Central Valley patterns with one or the other coastal region. In other words, if there are indeed differences between coastal Californian and inland Californian speech, they do not map to robust differences in speaker evaluation. This result mirrors the pilot study, in which Bay Area speakers' ratings patterned more closely with Lower Central Valley and Far Northern California speakers' ratings than with Southern California speakers' ratings.

The two self-comparison scales, *familiar* and *sounds like me*, were the only scales that yielded just one credible effect apiece, with Caucasians rated higher than Latinas/os on both scales. Interestingly, this difference was replicated across listener ethnicities; both Latina/o and Caucasian listeners gave higher *familiar* and *sounds like me* ratings to Caucasian speakers than Latina/o speakers. The scale *confident* was the only scale that yielded credible differences for all four primary predictors, suggesting a possible presence of interaction effects between these predictors. The following section addresses the interactions of primary predictors directly.

4.2.4. Interactions of primary predictors

This section addresses the second-order interactions of primary predictors: guise × region, guise × gender, guise × ethnicity, region × gender, region × ethnicity, and gender × ethnicity. Statements of credibility of interaction effects are based on the same decision rule for main effects—effects are deemed to be credible if the 95% HDI of the posterior contrast excludes zero—but the posterior contrast is computed differently for interaction effects than for main effects. For interaction effects, the differences at each step in the chain are between sums of interaction levels; if factor 1 has levels A & B and factor 2 has levels C & D, then the difference [(A × C) + (B × D)] – [(A × D) + (B × C)] is computed at each step in the chain. If the 95% HDI of the resulting posterior interaction contrast excludes zero, then the interaction can be interpreted two ways: the difference in ratings by factor 1 is credibly affected by the level of factor 2, or the difference in ratings by factor 2 is credibly affected by the level of factor 1.

As Figure 4.16 indicates, there were no credible effects of the guise × region interaction. That is, the difference between conservative and Californian guises applied equally to speakers from all three regions.

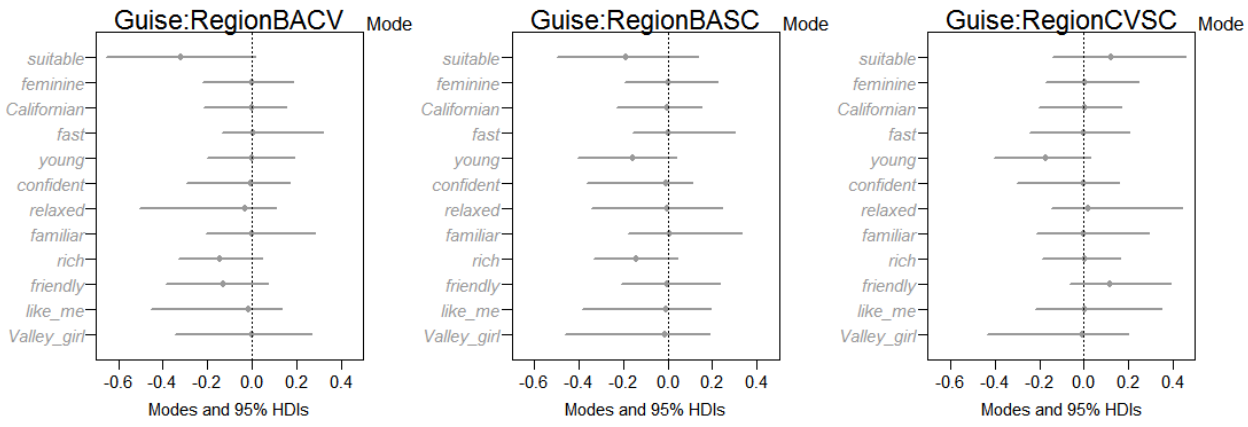


Figure 4.16. Modes and 95% HDIs of pairwise posterior scales interaction contrasts by guise and region.

Figure 4.17 displays interactions between guise and speaker gender, which yielded only one credible difference: *confident*. In particular, as Figure 4.18 indicates, whereas female speakers were rated equally *confident* in both guises, male speakers sounded more *confident* in the Californian guise than the conservative guise. Otherwise, there were no credible guise × gender interaction effects. That is, not only was the Californian guise rated more *Californian* and *Valley girl* than the conservative guise (4.2.1. Guise), this difference by guise was virtually the same for both female and male speakers. Likewise, the credible gender differences on *suitable*, *feminine*, *Californian*, *fast*, *relaxed*, *friendly*, and *Valley girl* (4.2.2. Speaker region, gender, and ethnicity) were virtually the same for both conservative and Californian guises. Whereas California-shifted speakers were rated higher on *Californian* and *Valley girl* than conservative speakers and females were rated higher than males, these effects did not compound one another.

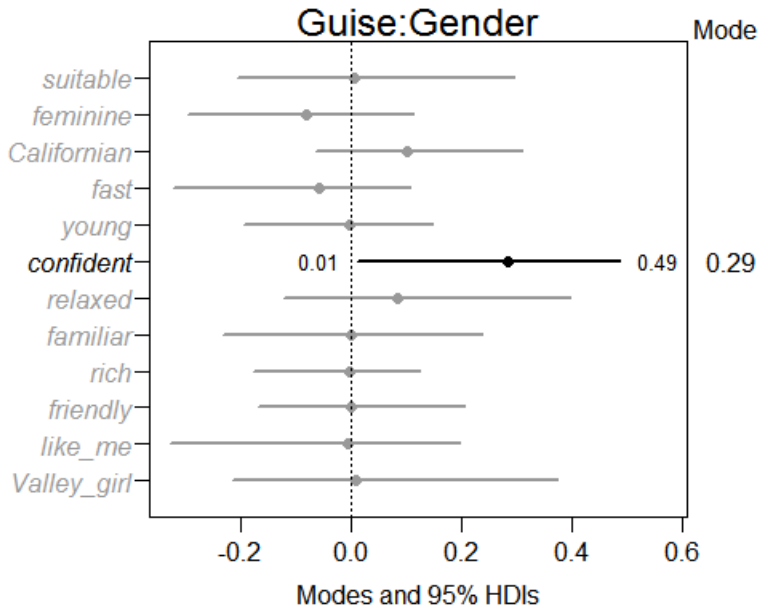


Figure 4.17. Modes and 95% HDIs of posterior scales interaction contrasts by guise and gender.

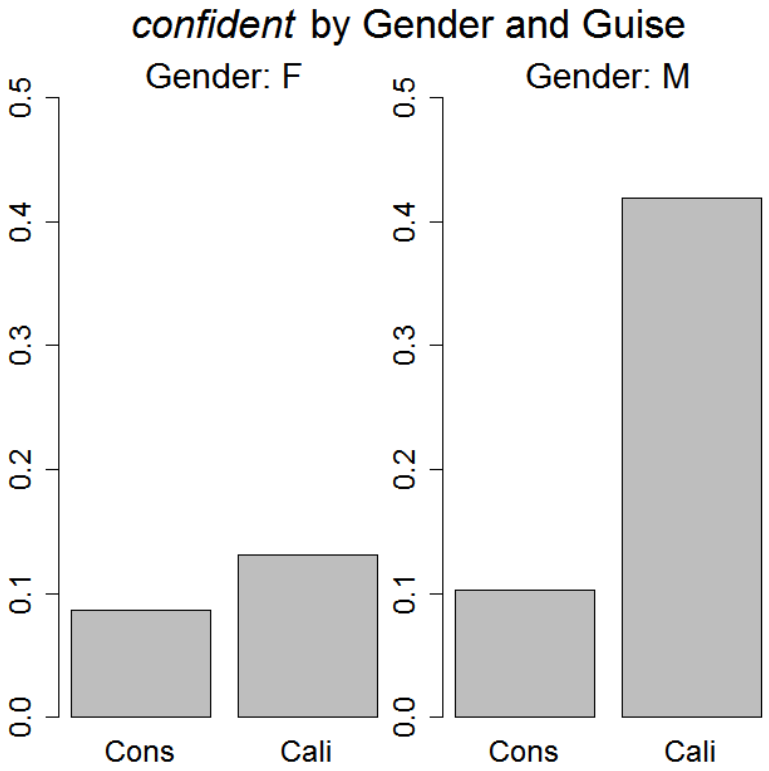


Figure 4.18. Mean standardized *confident* ratings by gender and guise.

As Figure 4.19 indicates, there were no credible effects of the guise × ethnicity interaction. That is, the difference in scale ratings between guises applied equally to both Latina/o and Caucasian speakers, regardless of whether guise or ethnicity had a credible main effect on ratings. This lack of interaction effect is evident in Figure 4.20, which displays mean *friendly* ratings by guise and ethnicity; here, the small conservative vs. Californian difference is the same for both ethnicities.

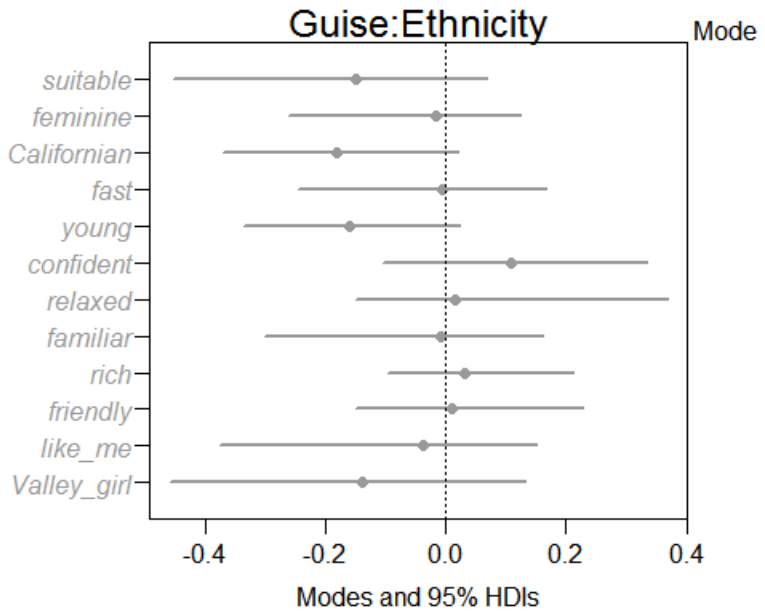


Figure 4.19. Modes and 95% HDIs of posterior scales interaction contrasts by guise and ethnicity.

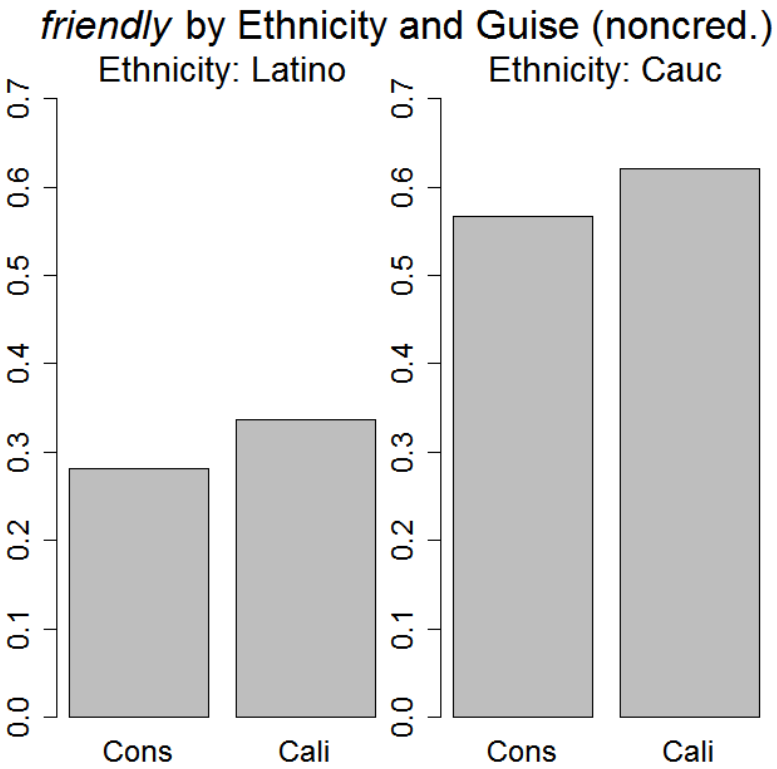


Figure 4.20. Mean standardized *friendly* ratings by ethnicity and guise (noncredible interaction effect).

The interaction effects of guise with other primary predictors represent additional evidence that guise registers weakly in Californian listeners' responses to language variation;

save for the interaction effect with gender on *confident*, the differences between listeners' reactions to the Californian guise and listeners' reactions to the conservative guise were much the same regardless of other speaker characteristics. However small these effects are, though, the lack of credibly nonzero interactions involving guise also indicates that the main effects of guise are a function of listeners' attending to these vocalic differences, not merely an artifact of other vocal characteristics present in the stimuli.

The pairwise interaction contrasts between region and gender, which are displayed in Figure 4.21, yielded at least one credible difference on all 12 scales; notably, all 12 scales yielded a credible interaction effect between gender and the Lower Central Valley vs. Southern California contrast. The two scales with the greatest differences were *sounds like a Valley girl* and *feminine*, for which region × gender means are displayed in Figures 4.22 and 4.23, respectively. Consistent with the main effect of gender on *Valley girl*, Bay Area females were rated higher than Bay Area males, and likewise for Lower Central Valley speakers; among Southern California speakers, however, *male* speakers were rated higher on *Valley girl* than female speakers. Unlike *Valley girl*, the region × gender means for *feminine* in Figure 4.23 indicate that the gender contrasts differed interregionally by magnitude, not direction; whereas Lower Central Valley females were rated more *feminine* than Lower Central Valley males by over 2.5 standard deviations, the contrast for Bay Area speakers was 2 standard deviations and for Southern California speakers 1.5 standard deviations.

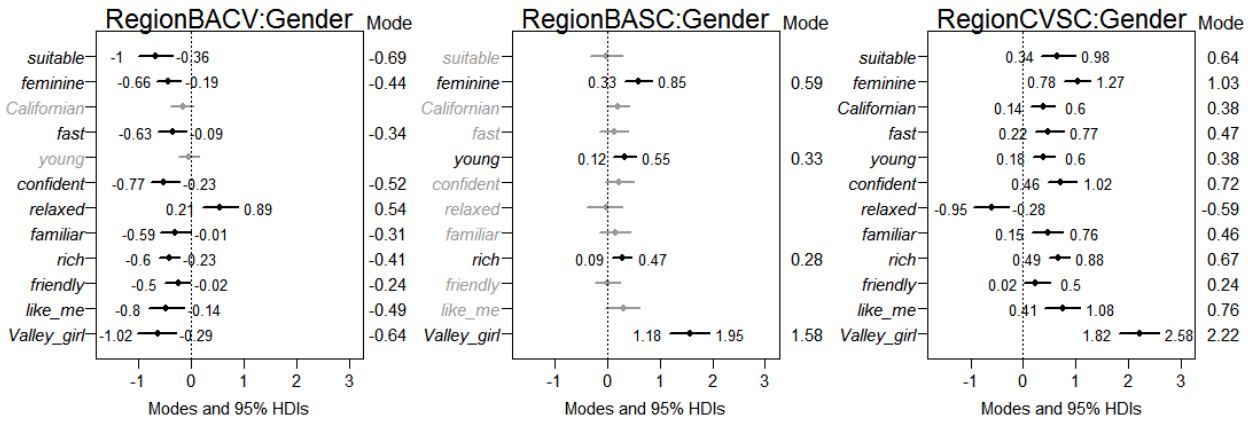


Figure 4.21. Modes and 95% HDIs of pairwise posterior scales interaction contrasts by region and gender.

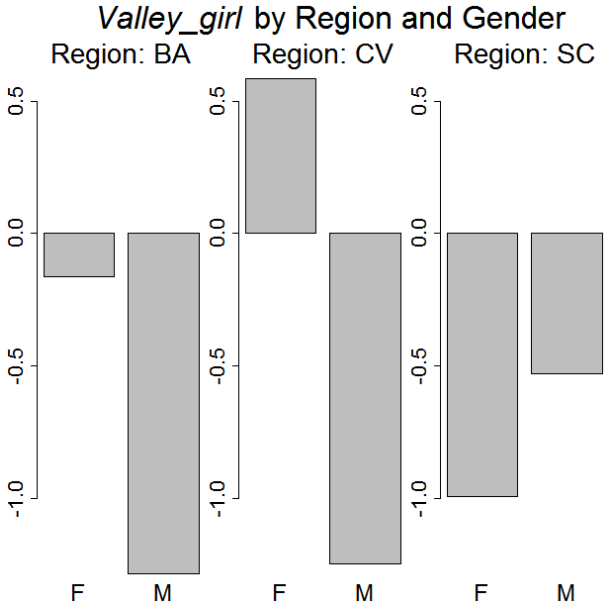


Figure 4.22. Mean standardized *sounds like a Valley girl* ratings by region and gender.

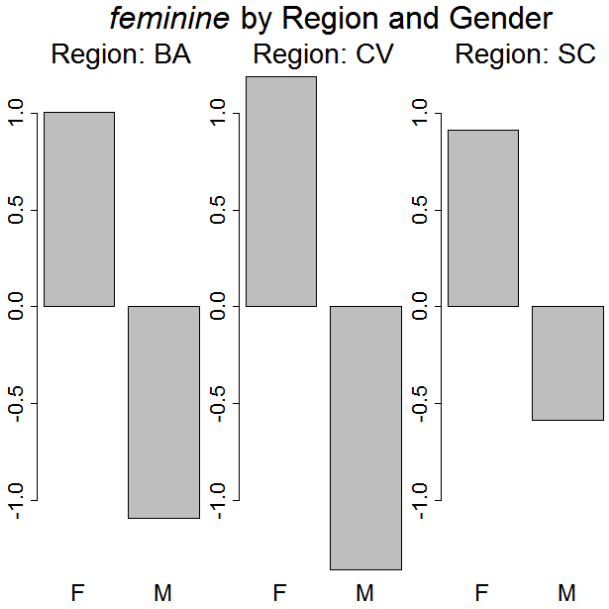


Figure 4.23. Mean standardized *feminine* ratings by region and gender.

Taken together, the region × gender results for *Valley girl* and *feminine* are useful not for what they indicate about speaker regions themselves—as mentioned in the previous subsection, the results involving speaker region are best taken with a grain of salt—but because they demonstrate that, among Californians, sounding like a Valley girl is an attitudinal construct that is similar to but distinct from merely sounding feminine. That is, the fact that the male Southern California speakers in this study apparently sounded more like Valley girls than their female counterparts is likely not generalizable to all Southern California male and female speakers;

however, the fact that listeners rated the female Southern California speakers more feminine than their male counterparts despite sounding less like Valley girls indicates that Californians hear Valley girls as representing something beyond mere femininity.

The pairwise interactions between region and ethnicity, as displayed in Figure 4.24, indicate six credible effects, with one credible difference on one scale, *confident*, and two credible differences on five scales, *suitable*, *feminine*, *familiar*, *friendly*, and *sounds like a Valley girl*.

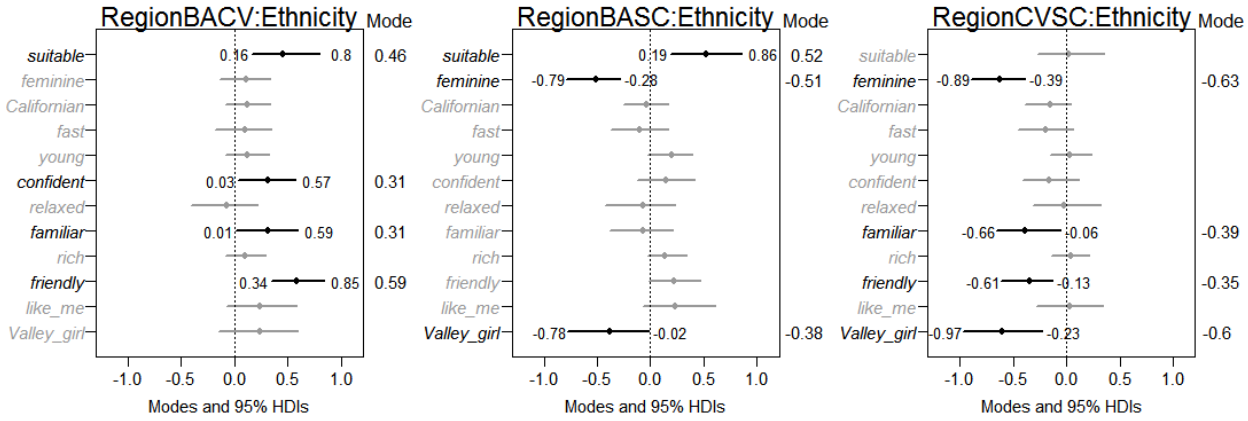


Figure 4.24. Modes and 95% HDIs of pairwise posterior scales interaction contrasts by region and ethnicity.

The interaction between gender and ethnicity, as displayed in Figure 4.25, yielded five credible differences: *feminine*, *fast*, *confident*, *relaxed*, and *friendly*. Each of these scales also had independent credible main effects for both gender and ethnicity, with female speakers patterning with Latina/o speakers and male speakers patterning with Caucasian speakers in each case (e.g., females were rated credibly higher on *relaxed* regardless of ethnicity, as were Latinas/os regardless of gender; males were rated credibly higher on *fast* regardless of ethnicity, as were Caucasians regardless of gender). The gender × ethnicity means displayed in Figures 4.26 and 4.27 elucidate this pattern: on each of these scales (including the three not pictured in Figures

4.26 and 4.27), male Caucasian speakers were rated substantially differently than female Latinas, female Caucasians, or male Latinos.

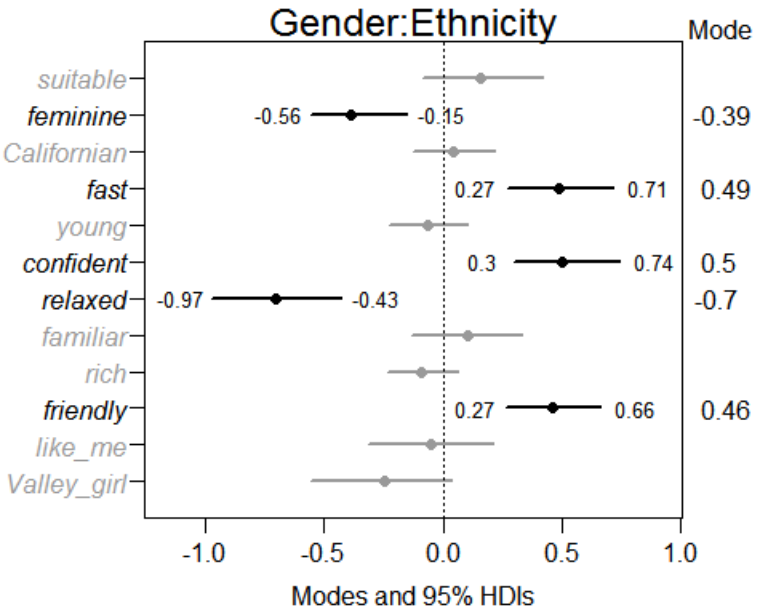


Figure 4.25. Modes and 95% HDIs of posterior scales interaction contrasts by gender and ethnicity.

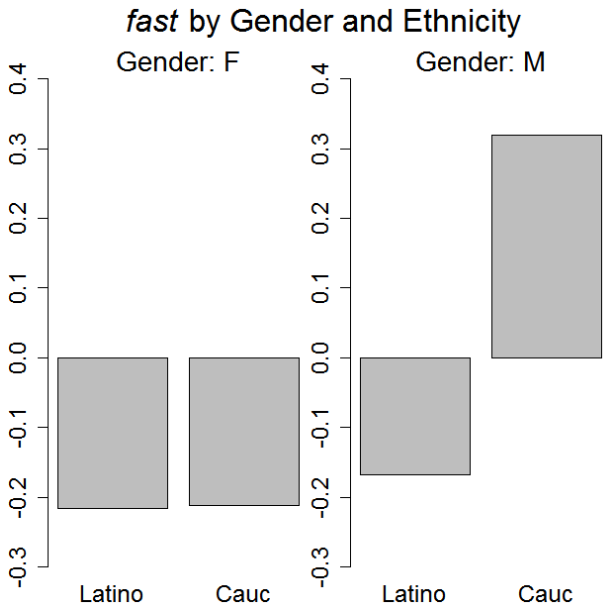


Figure 4.26. Mean standardized *fast* ratings by gender and ethnicity.

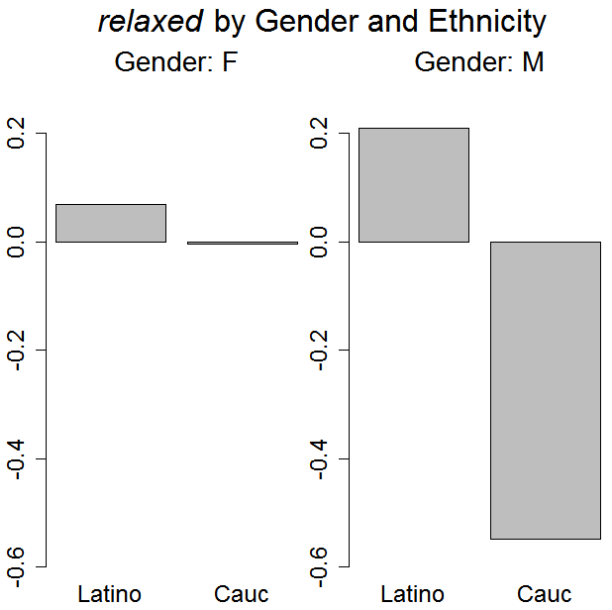


Figure 4.27. Mean standardized *relaxed* ratings by gender and ethnicity.

To summarize, guise yielded only one credible interaction effect with primary predictors, as male speakers in the Californian guise received a substantial boost in *confident* ratings relative to the conservative guise, whereas females were rated equally low on *confident* regardless of guise; in other words, the CVS makes males, but not females, sound more confident. The general lack of interactions involving guise reinforces two results emerging from the analysis of guise as a main effect: the effect of guise is small in magnitude, registering less strongly than other speaker characteristics on listeners' ratings—but it is real, as the main effects of guise are not simply a function of guise piggybacking on other speaker characteristics. The interactions not involving guise, by contrast, indicate a higher degree of interrelation between region, gender, and ethnicity, with the region × gender interaction yielding credible effects on all 12 scales. The gender × ethnicity interactions showed that male Caucasian speakers stand apart from other speakers on *feminine*, *fast*, *confident*, *relaxed*, and *friendly*, suggesting that maleness and Caucasian-ness may compound the effects of one another in the ears of Californian listeners.

Finally, the region × gender interaction showed that the percept of sounding like a Valley girl is influenced by, but not wholly tied up in, the percept of femininity.

4.2.5. Listener region and perceived region

As Figure 4.28 indicates, listener region accounted for few credible differences. Lower Central Valley listeners rated speakers more *confident* overall than did Southern California listeners, and less *young* overall than did both Bay Area and Southern California listeners. There were no credible differences in evaluation between listeners from the two coastal regions.

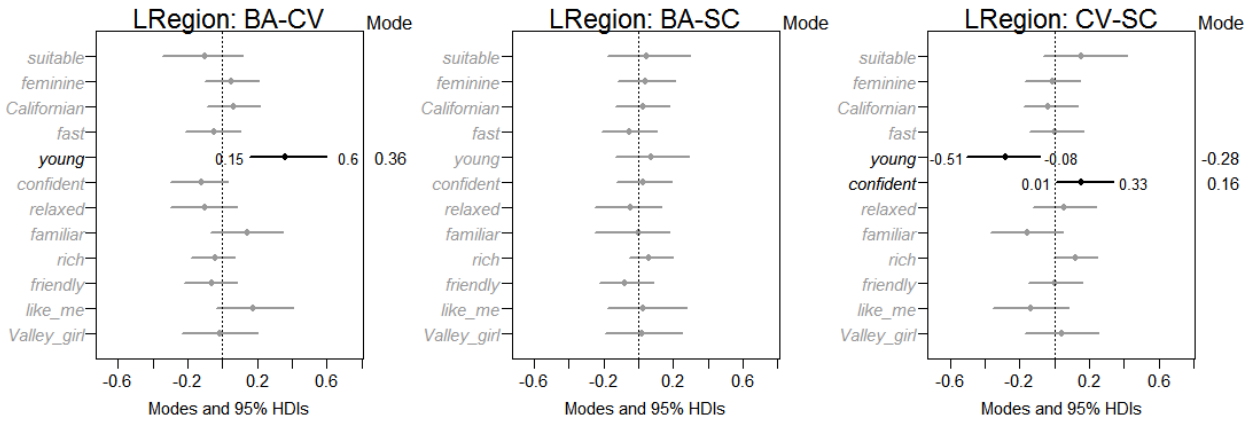


Figure 4.28. Modes and 95% HDIs of pairwise posterior scales contrasts by listener region (LRegion).

Figure 4.29 displays pairwise differences for perceived region, the region that listeners thought that speakers were from, with a summary of credible differences in Table 4.3.[12] Of course, listeners did not perform above chance in accurately recognizing speakers' region (see 4.1. Regional identification, above). It is nevertheless useful to assess how perceived region affected speaker evaluation as a means of assessing folk-linguistic notions of speech from each region. I will first assess differences among perceived Californian regions (the top row of

[12] Trials were categorized into perceived regions based on the regional identification response for each trial. This is a different procedure than was used in the pilot study, in which speakers were categorized into perceived regions based on the region into which a plurality of listeners placed a speaker. Although these results are really about perceptions within each *trial*, for brevity's sake I discuss these results in terms of where *speakers* were perceived to be from.

Figure 4.29), then differences between speakers perceived to be from California vs. outside California (the bottom row of Figure 4.29).

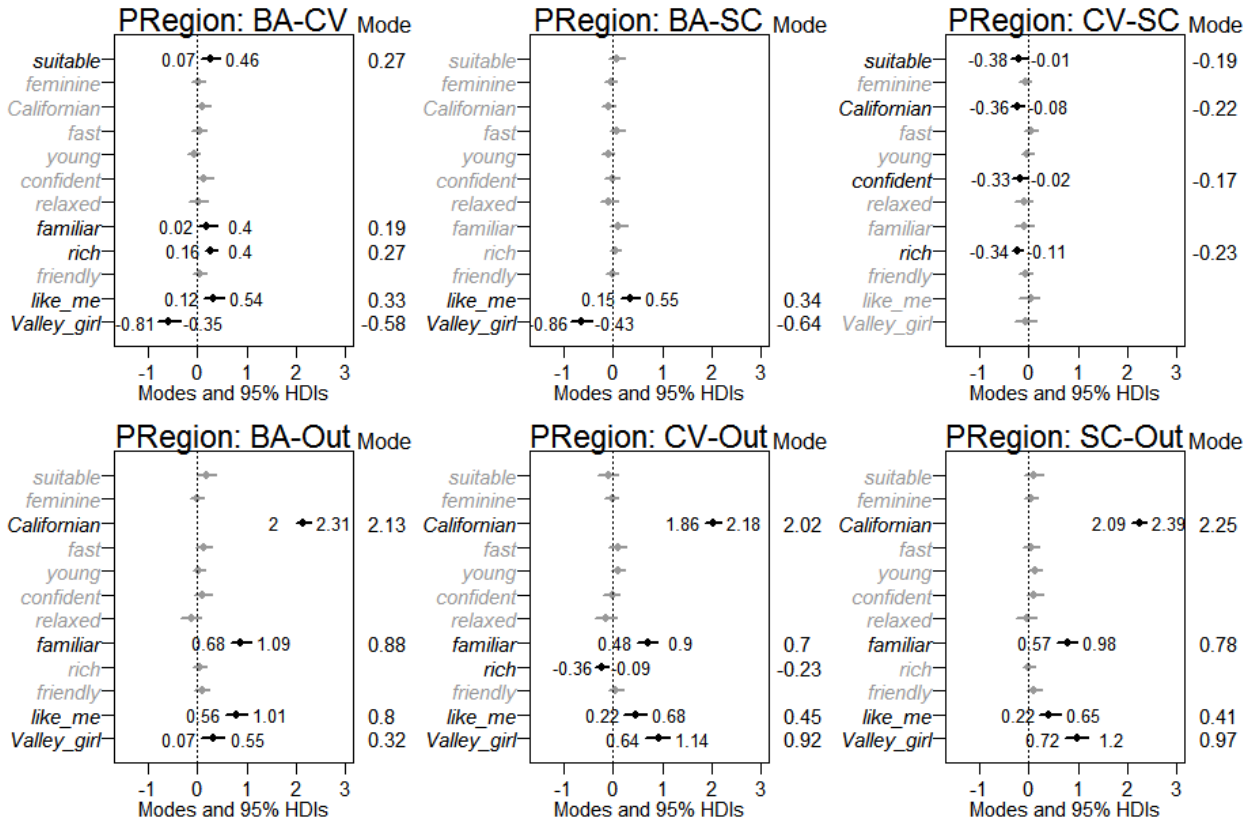


Figure 4.29. Modes and 95% HDIs of pairwise posterior scales contrasts by perceived region (PRegion). Out indicates outside California.

Scale	Perceived region	Scale	Perceived region
suitable	BA, SC > CV	*relaxed*	
feminine		*familiar*	BA (SC) > CV (SC) > Out
Californian	SC (BA) > CV (BA) > Out	*rich*	BA, SC, Out > CV
fast		*friendly*	
young		*like me*	BA > SC, CV > Out
confident	SC > CV	*Valley girl*	CV, SC > BA > Out

Table 4.3. Credible differences by scale for perceived region.
The notation X > Y indicates that X was rated credibly higher on that scale than Y. The notation X, Y > Z indicates that X and Y did not receive credibly different ratings, but both X and Y were rated credibly higher than Z. The notation X (Z) > Y (Z) > W indicates that Z was rated credibly higher than W, but not credibly different than X or Y. Blank cells indicate no credible difference.

The top row of Figure 4.29 indicates that seven scales yielded at least one credible intra-Californian pairwise difference: *suitable, Californian, confident, familiar, rich, sounds like me,* and *sounds like a Valley girl*. Speakers perceived to be from the Bay Area were rated more *familiar* than speakers perceived to be from the Lower Central Valley, more *like me* than speakers perceived to be from either the Lower Central Valley or Southern California, and less *Valley girl* than speakers perceived to be from either the Lower Central Valley or Southern California. Speakers perceived to be from Southern California were rated more *Californian* and *confident* than speakers perceived to be from the Lower Central Valley. Speakers perceived to be from either coastal region were rated more *suitable* and *rich* than speakers perceived to be from the Lower Central Valley.

Among Californian regions, perceived region had the greatest effect on *sounds like a Valley girl*, as speakers perceived to be from the Lower Central Valley and especially Southern California were rated more *Valley girl* than speakers perceived to be from the Bay Area; in other words, Bay Area speech is excluded from qualities associated with the notion of sounding like a Valley girl. The set of credible differences on *Valley girl* was not the same as that for *Californian*, for which speakers perceived to be from Southern California were rated higher than

the Lower Central Valley, or *feminine*, for which no differences by perceived region emerged as credible. The mean *Californian* and *Valley girl* ratings by perceived region, which are displayed in Figure 4.30, demonstrate another difference between these two scales; listeners generally accepted the notion that speakers sounded Californian (except where they identified speakers as from outside California) and generally rejected the notion that speakers sounded like Valley girls, but both scales yielded ratings that differentiated perceived Californian regions. (This result echoes the mean *Californian* and *Valley girl* ratings by guise in Figure 4.12.) In sum, the results by perceived region add to the evidence that Californian listeners construct the notion of sounding like a Valley girl as related to but distinct from 'Californian-ness' and femininity.

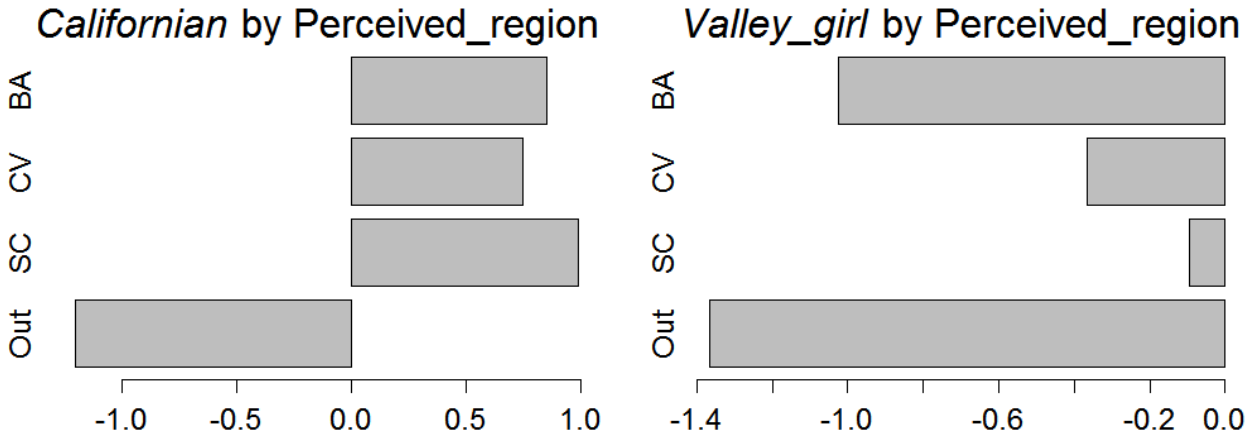


Figure 4.30. Mean standardized *Californian* and *sounds like a Valley girl* ratings by perceived region.

It is also noteworthy that speakers perceived to be from the Lower Central Valley received the lowest ratings for almost all of the scales that yielded a credible intra-Californian perceived-region difference (all save for *Valley girl*), which echoes the general stigmatization of the Lower Central Valley found in the pilot (see 2.2.3. Interaction of identification and language attitudes, above). Unlike the pilot, however, not only were Lower Central Valley listeners represented in the main study listener sample, they actually represented the largest group of listeners in this sample. In addition, speakers perceived to be from the Bay Area and Southern

California clustered together on *suitable* and *rich*, the two scales that most closely align with "status" traits.

The bottom row of Figure 4.29 displays pairwise contrasts between speakers perceived to be from each Californian region and speakers perceived to be from outside California (i.e., interstate contrasts). Notably, all of the scales with credible interstate differences also yielded at least one credible *intra*state difference; *Californian*, *familiar*, *rich*, *sounds like me*, and *sounds like a Valley girl* yielded both intrastate differences and interstate differences, whereas *suitable* and *confident* yielded only intrastate differences. The effects of the interstate differences tended to be stronger than those of the intrastate differences (e.g., for *familiar*, the pairwise posterior contrast between speakers believed to be from the Bay Area vs. the Lower Central Valley had a mode of 0.19, whereas the contrast between speakers believed to be from the Bay Area vs. outside California had a mode of 0.88). These interstate differences indicate, in other words, a substantial separation in attitudes toward being from California vs. elsewhere. Speakers perceived to be from the Bay Area, Lower Central Valley, or Southern California were rated more *like me*, *familiar*, *Valley girl*, and—by over 2 standard deviations—*Californian* than speakers perceived to be from outside California. Of course, the difference in *Californian* ratings is to be expected, as the same listeners also identified the speaker as being from within or outside California, but the fact that these interstate differences also mapped to *like me*, *familiar*, and *Valley girl* indicates an association between these traits and listeners' beliefs about how Californian speakers sound.

Speakers perceived to be from outside California were also rated more *rich* than speakers perceived to be from the Lower Central Valley. This credible difference was the only one for which speakers perceived to be from outside California were rated higher than speakers

perceived to be from any Californian region, a result that aligns with the pattern of stigmatization of Lower Central Valley speech in the scales data. Unlike the pilot study, which found a general stigmatization of non-Californian speech, the credible differences found in the main study between speakers perceived to be from California vs. outside California do not obviously reflect a stigmatization of non-Californian speech—putative non-Californians were not evaluated as slower, less confident, or less suitable for a speaking job. By contrast, the main study's Californian listeners' assignment of the Valley Girl label to themselves may represent a degree of *self*-stigmatization (relative to a lower baseline, of course, as indicated by Figure 4.30)—but only if we assume that the "Valley girl" label and character type carry the same stigma to Californians as is suggested by the original conception of the stereotype (see 1.6.3. California enregisterment, above).

In summary, listener region had little effect on listeners' scale ratings; the exceptions were three contrasts involving Lower Central Valley listeners, who rated speakers more *confident* overall than did Southern California listeners, and who rated speakers less *young* overall than did both Bay Area and Southern California listeners. Perceived region had a much greater effect on listeners' scale ratings, with credible intrastate differences on *suitable, Californian, confident, familiar, rich, sounds like me,* and *sounds like a Valley girl*; the intrastate results for perceived region indicate that these regions have a folk-linguistic reality among Californian listeners. These results also add to the evidence that Californian listeners construct the notion of sounding like a Valley girl as related to but distinct from Californian-ness and femininity. The contrasts between speakers perceived to be from outside California and those perceived to be from California revealed strong effects on *Californian, familiar, like me,* and *Valley girl,* indicating that Californians set themselves apart from non-Californians to a greater

degree than they attend to intrastate folk-linguistic differences. Unlike the pilot study, however, these interstate differences in perceived region did not lead to the stigmatization of putative non-Californian speakers. By contrast, the main study results for perceived region reveal a general stigmatization of the Lower Central Valley, even as compared with non-Californians.

4.2.6. Interactions of guise with listener region and perceived region

The guise × listener region interaction was analyzed to address whether the social meanings of the California Vowel Shift differ by listener region; that is, did listener region affect conservative vs. Californian rating differences? Figure 4.31 reveals only one credible interaction: an interaction on *feminine* between guise and the Bay Area vs. Southern California pairwise contrast. As Figure 4.32 indicates, this interaction is a result of complementary attitudes among listeners from these regions; the Californian guise was rated less *feminine* by Bay Area listeners and more *feminine* by Southern California listeners. That is, whereas Bay Area listeners associated the CVS with masculinity, Southern California listeners associated the CVS with femininity. Aside from *feminine*, however, no other scales yielded credible interaction contrasts. This means that, for example, the association of California-shifted vowels with higher *Californian* ratings was represented more or less evenly for each listener region (as evidenced in Figure 4.33); likewise, the main effect of listener region on *young* ratings was not credibly affected by guise (as evidenced in Figure 4.34).

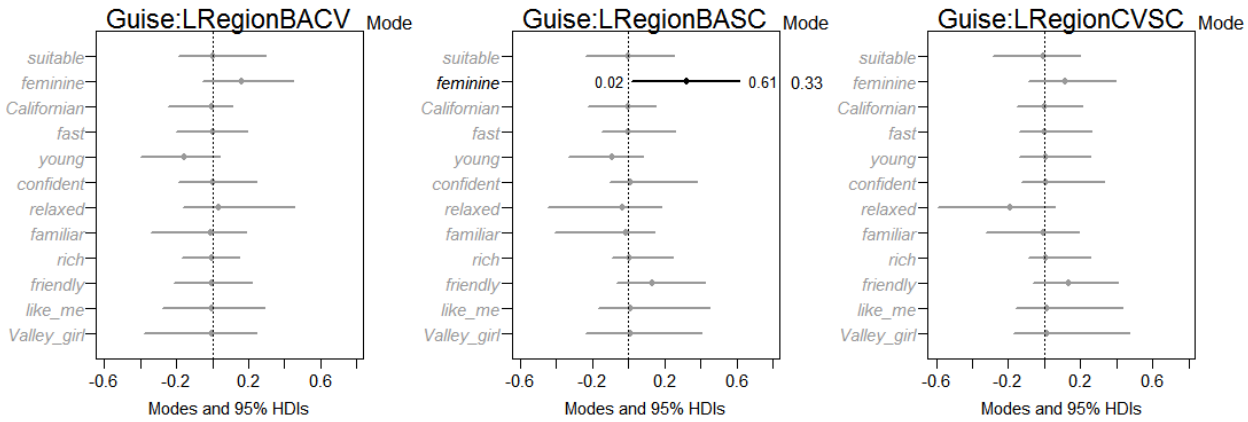


Figure 4.31. Modes and 95% HDIs of pairwise posterior scales interaction contrasts by guise and listener region.

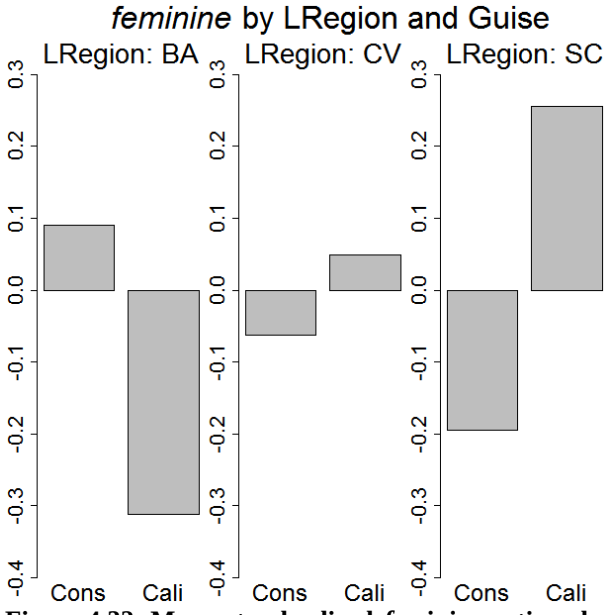


Figure 4.32. Mean standardized *feminine* ratings by listener region and guise.

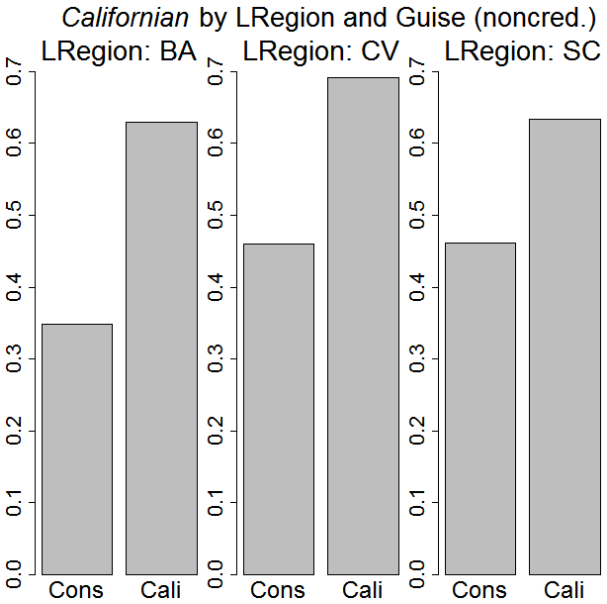


Figure 4.33. Mean standardized *Californian* ratings by listener region and guise (noncredible interaction effects).

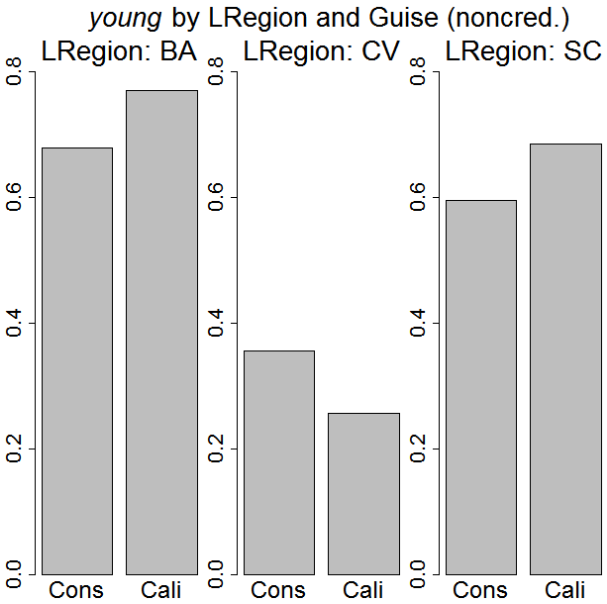


Figure 4.34. Mean standardized *young* ratings by listener region and guise (noncredible interaction effects).

The analysis of guise × perceived region interaction effects addresses whether the social meanings of the CVS differ by perceived region; for example, was the difference in *Valley girl* ratings larger if speakers were thought to be from Southern California than if they were thought to be from the Bay Area? As Figure 4.35 indicates, there were no credible pairwise guise × perceived region interaction differences; in fact, over two thirds of all posterior interaction contrast modes were within 0.01 standard deviations of zero. These results indicate rather strongly that the effect of guise is not sensitive to perceived region, and vice versa.

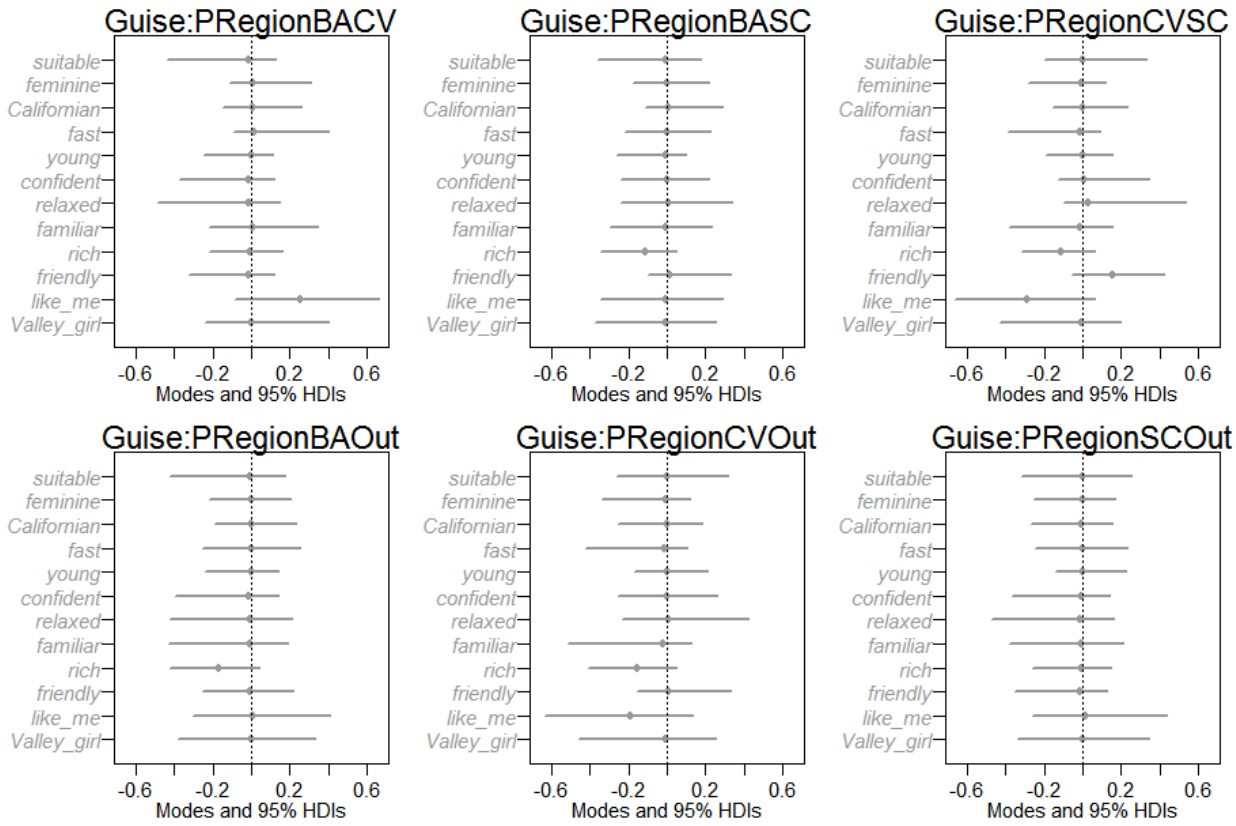


Figure 4.35. Modes and 95% HDIs of pairwise posterior scales interaction contrasts by guise and perceived region.

To summarize, the interactions of guise with listener region and perceived region yielded just one credible interaction effect, as Bay Area listeners associated the CVS with masculinity whereas Southern California listeners associated the CVS with femininity. No guise × perceived region interactions were credibly nonzero—most were rather close to zero—suggesting that the perceptual effects of the CVS on the one hand and folk-linguistic beliefs about speakers from certain regions on the other operate independently of one another. These results also mirror the interactions of guise with speaker region, speaker gender, and speaker ethnicity (see 4.2.4. Interactions of primary predictors, above), which yielded just one credible interaction effect (guise × gender on *confident*). Together, the analyses of interactions involving guise suggest that the CVS affects listeners' perceptions in largely the same way regardless of other speaker (or

even listener) characteristics; moreover, these non-credible interaction effects lend more credence to the main effects of guise—an association between the CVS and sounding Californian, confident, and like a Valley girl—as these effects are not likely to be the result of other vocal characteristics present in the stimuli.

4.2.7. Interaction of listener region with perceived region

The interaction of listener region × perceived region included 18 pairwise posterior interaction contrasts. (Graphs of all 18 contrasts are omitted here for the sake of space.) These pairwise contrasts yielded at least one credible difference on five scales: *Californian*, *fast*, *familiar*, *friendly*, and *sounds like me*. *Fast* yielded credible differences for 3 of the 18 pairwise interaction contrasts, and *friendly* yielded credibly differences for 4 of the 18 pairwise interaction contrasts; *Californian*, *familiar*, and *like me* all yielded 8 credible differences. A comparison of the three contrasts for which the pair of listener regions matches the pair of perceived regions, displayed in Figure 4.36, addresses how listeners rated speakers they perceived to be from the same region as themselves differently from speakers they perceived to be from other regions—in short, whether claiming effects emerge for regional ingroup perceivers (see 1.4.3. Dialect recognition and 2.2.3. Interaction of identification and language attitudes, above).

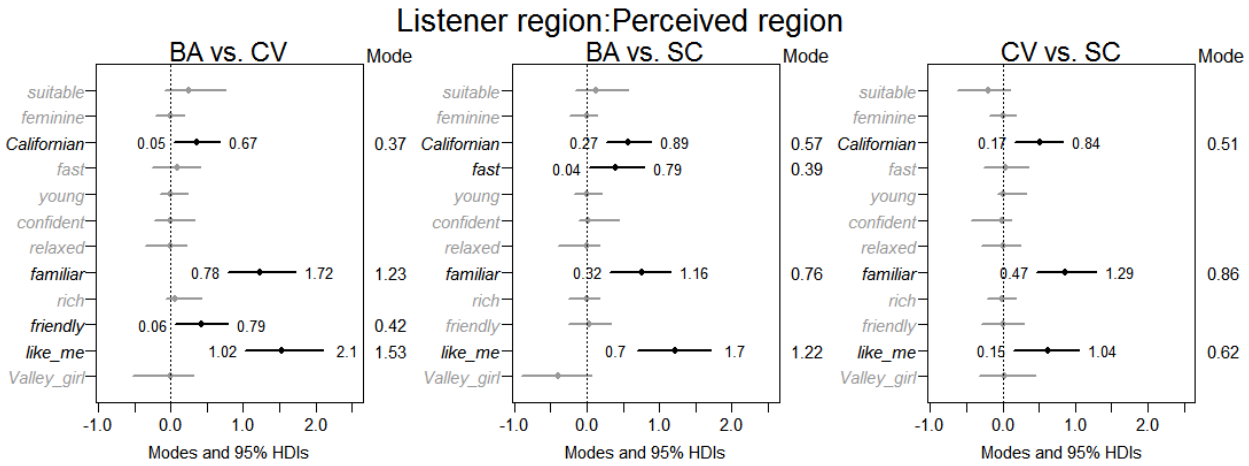


Figure 4.36. Modes and 95% HDIs of the three pairwise posterior scales interaction contrasts by listener region and perceived region for which the pair of listener regions matches the pair of perceived regions.

Californian, *familiar*, and *like me* figured into all three of these contrasts, with similar effects: listeners gave higher ratings on all three of these scales to perceived ingroup members. This ingroup preference was especially strong for *like me*, and least strong for *Californian*. The magnitudes of these listener region–perceived region matched-pair interaction contrasts provides additional evidence for the strength of these claiming effects. In other words, these ratings represent true claiming effects on *Californian*, *familiar*, and *like me*, as listeners generally ascribed the highest ratings on these scales to speakers they believed to be part of their regional ingroup. As Figure 4.37 shows, this effect is weaker for *Californian* among Lower Central Valley listeners, who give a slight preference on Californian-ness to speakers they believed to be from Southern California.

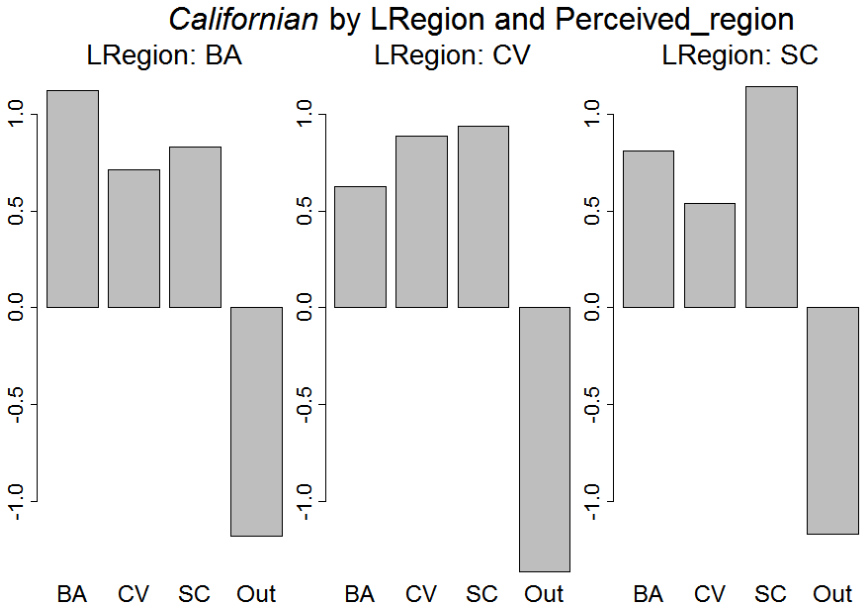


Figure 4.37. Mean standardized *Californian* ratings by listener region and perceived region.

The results in Figure 4.36 also suggest partial evidence for two claiming effects for Bay Area listeners beyond *Californian*, *familiar*, and *sounds like me*: *fast* and *friendly*. On *fast*, the difference between Bay Area listeners' ratings for speakers they believed to be from the Bay Area vs. Southern California was credibly larger than the same perceived-region contrast for Southern California listeners; similarly, Bay Area listeners gave speakers they perceived to be from the Bay Area a credibly larger boost on *friendly* over speakers they perceived to be from the Lower Central Valley, relative to the same perceived-region contrast for Lower Central Valley listeners. These results suggest some evidence that Bay Area listeners claim the speakers who they hear as fast-speaking and friendly, although the strength of this evidence is mitigated by the fact that for both scales, only one of the two interaction contrasts involving the Bay Area is credibly nonzero.

Interestingly, the two self-comparison items did not exhibit the same polarity; the mean ratings on *familiar* were generally positive (as evidenced in Figure 4.38), whereas the mean

ratings on *like me* were almost all negative (as evidenced in Figure 4.39). Nevertheless, listeners from each region gave their highest *familiar* ratings to perceived regional ingroup members, and Bay Area and Lower Central Valley listeners assigned their least negative *like me* ratings to perceived regional ingroup members; on *like me*, Southern California listeners slightly favored speakers they perceived to be from the Bay Area. In other words, these listeners perceived speakers in general as fairly familiar, but especially so if they believed speakers belonged to their regional ingroup. Likewise, these listeners did not perceive speakers in general as sounding much like them, but this distancing effect was mitigated for the speakers who they believed belonged to their regional ingroup. This effect is reminiscent of the effect of guise on the *Californian* and *sounds like a Valley girl* ratings displayed in Figures 4.12 and 4.30; a listener's perception that a speaker belongs to the listener's regional ingroup boosts the speaker's *familiar* and *sounds like me* ratings, regardless of the higher baseline rating for *familiar* and the lower baseline rating for *like me*.

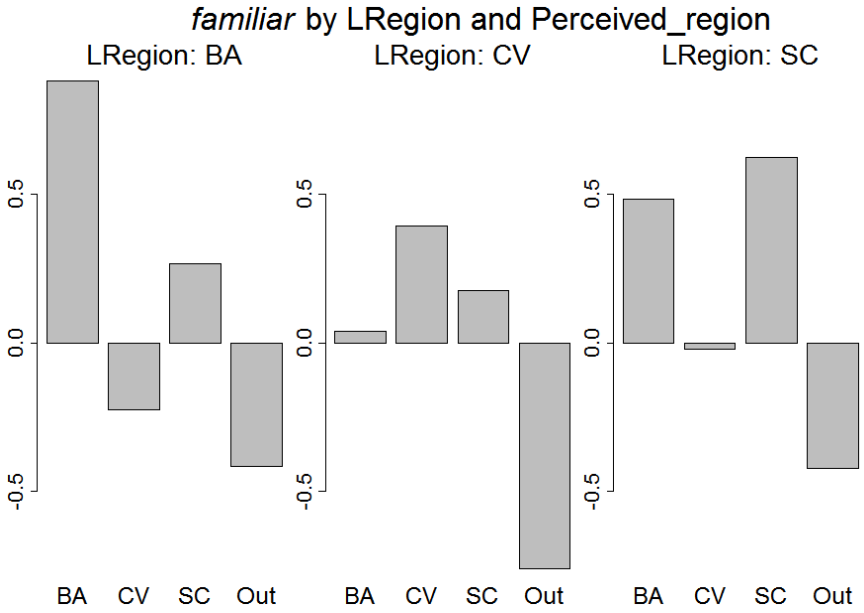


Figure 4.38. Mean standardized *familiar* ratings by listener region and perceived region.

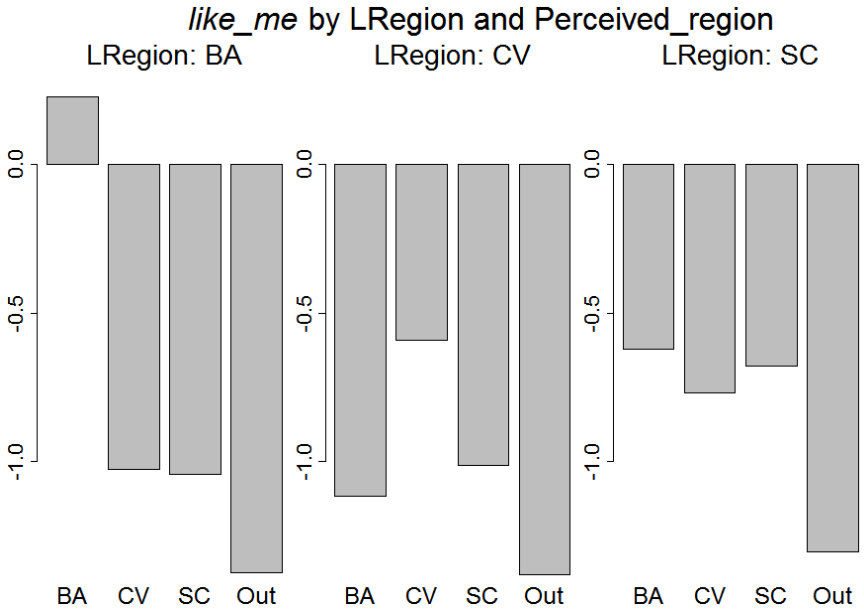


Figure 4.39. Mean standardized *sounds like me* ratings by listener region and perceived region.

The claiming effects found in the main study are similar to the pilot's findings about claiming effects among listeners from the Bay Area and Southern California (2.2.3. Interaction of identification and language attitudes). In the main study, as in the pilot, listeners from both

coastal regions claimed the speakers who they perceived as members of their regional ingroup on *Californian* and *speaks like me*. The pilot found that Southern California listeners, but not Bay Area listeners, claimed speakers who they rated highest on *familiar*; in the main study, this claiming effect for *familiar* extended to Bay Area listeners. The main study revealed that these claiming effects for *Californian*, *familiar*, and *sounds like me* were shared by Lower Central Valley listeners. The main study also provided evidence for the relative strength of these claiming effects, even against a low overall baseline for *sounds like me*. Finally, whereas the pilot suggested a claiming effect for Bay Area listeners on a "pleasantness" trait, *attractive*, the main study provided partial evidence for Bay Area listeners claiming both a "pleasantness" trait, *friendly*, and a "status" (or "correctness") trait, *fast*.

In addition, the stability of the claiming results from the pilot to the main study contrasts with the results for scale ratings by (actual) speaker region, which show substantial inconsistencies between the pilot and main study. I argued above (4.2.3. Summary: Primary predictors) that these inconsistencies were the result of idiosyncratic differences between the pilot and main study speakers who represented each speaker region, which suggests that speaker region is not a reliable basis for Californian listeners' ratings. The stability of the claiming results, by contrast, demonstrates that these regions have an *evaluative* reality beyond any actual differences in speech; that is, folk-linguistic images of these regions exist in Californian listeners' heads, even if these listeners do not consistently rate speakers by their actual region—or cannot reliably differentiate speakers by region.

In sum, the analysis of listener region × perceived region interactions revealed several strong claiming effects among Californian listeners, most notably for *Californian*, *familiar*, and *sounds like me*. These claiming effects were evident among all three listener regions, including

the Lower Central Valley, in spite of the general stigmatization of the Lower Central Valley (including by Lower Central Valley listeners). In other words, Californian listeners across the state believe speakers from their region to sound the most Californian, the most familiar, and the most like them.

4.2.8. Non-controlled predictors

As mentioned above, non-controlled predictors were analyzed via individual two-predictor models for each non-controlled predictor and each scale; speaker was included as the second predictor in these models as an un-analyzed 'random effect' (fn. 6).

Five non-controlled predictors had no credible effect on scales responses: listener gender, listener ethnicity, listener mono/bi/multilingualism, listener experience with linguistics, and student status. Five non-controlled predictors yielded at least one credible difference on scales responses: trial order (ten credible differences), number of listens (seven credible differences), listener mobility (seven), listener age (two), and listening device (one). Notably, the non-controlled predictors with at least one credible difference also had a greater number of factor levels (and thus a greater number of pairwise contrasts between factor levels).[13] In addition, *suitable* and/or *young* were involved in all five of the non-controlled predictors with at least one credible difference. In general, the results involving even the non-controlled predictors that yielded credible differences indicate that non-controlled predictors had little effect on scale ratings overall.

Trial order yielded 10 credible differences (with 15 pairwise contrasts for each of the 12 scales). In general, these differences show that listeners rated speakers less *suitable* in trial 1 and

[13] This is less true of the non-controlled predictors that credibly affected regional identification responses, which included the two-level factors listener gender and student status (4.1.4. Non-controlled predictors).

more *young* in trials 1 and 2, but listeners moderated their ratings on both *suitable* and *young* beyond trial 2. No scales other than *suitable* and *young* (which yielded five credible differences apiece) were credibly affected by trial order, and nine of the credible differences (with the exception of trial 2 vs. trial 6 on *young*) involved a difference between trial 1 and a later trial. In short, aside from this effect of early trials on *suitable* and *young*, listeners' scale ratings did not systematically change over the course of the listening task.

Listeners self-reported the number of times they listened to each clip for each trial. The distribution of self-reported listens skewed high, with a median of two listens and a mean of 2.5 listens per trial; the counts of trials with more than five listens were in the single digits for each number of listens. Number of listens was analyzed as a categorical factor with five levels: one listen (28.1% of trials), two listens (34.1% of trials), three (20.2%), four (9.5%), and five or more listens (8.1%); with five levels, the analysis of number of listens involved 10 pairwise contrasts for each of the 12 scales. Number of listens yielded seven credible differences, with four involving *suitable,* two involving *fast,* and one involving *familiar*. The differences on *suitable* all involved a positive contrast between a lower vs. higher number of listens, suggesting that the more times that listeners heard a speaker, the less *suitable* they rated the speaker. The differences involving *fast* and *familiar* do not show as clear a pattern as those involving *suitable*. Two listens meant higher ratings on *fast* than one listen or three listens, which does not suggest a cumulative effect of more listens. Two listens also meant higher ratings on *familiar* than five or more listens, but this single difference fails to suggest a real trend. Aside from the clear effect on *suitable*, listeners' scale ratings were not affected by number of listens.

Listener mobility yielded seven credible differences (with 10 pairwise contrasts for each of the 12 scales): five differences on *young* and two on *suitable*. Neither set of credible pairwise

differences suggested a clear cumulative effect of increased mobility on listeners' ratings. The two credible differences on *suitable*, for example, showed opposite effects of increased mobility on listeners' ratings; listeners who had lived in the same city their entire lives rated speakers credibly less *suitable* than did listeners who had lived in multiple regions of California, but the latter group of listeners rated speakers credibly *more suitable* than did listeners who had lived in multiple countries. Similarly, among the five pairwise differences on *young*, two involved higher ratings from the more mobile category of listeners and three involved higher ratings from the less mobile category of listeners. Rather than indicating any cumulative effect of listener mobility on *young* ratings, is likely that these differences reflect the preponderance of Lower Central Valley listeners, who rated speakers across the board less *young* than did Bay Area or Southern California listeners (see 4.2.3. Summary: Primary predictors, above), in the lower categories of mobility, especially the least mobile category: listeners who had lived in the same city their entire lives.[14] Once this effect on *young* is accounted for as an indirect effect of listener region, there are no clear effects of listener mobility on scale ratings.

Interestingly, the pattern of the effect of listener mobility on *young* mirrored that of the non-controlled predictors' effects on regional identification (4.1.4. Non-controlled predictors); the predictor levels favoring a Lower Central Valley identification were also those for which Lower Central Valley listeners were more heavily represented (females, Latinas/os, students, and listeners who had lived in the same city). In that sense, it is somewhat surprising that contrasts by listener gender, listener ethnicity, and listener student status did not also account for credible

[14] Of the 18 listeners in the main study sample who had lived in the same city their entire lives, 13 were from the Lower Central Valley. This effect is partially due to the fact that numerous Lower Central Valley listeners were recruited from Fresno State University, which is located in the most populous city in this region.

differences on *young* or other scales. It is possible that among Lower Central Valley listeners, regional identification is simply more salient than speaker attributes not tied to region.

As mentioned above (3.4.5. Listeners), the main study listener sample skewed young, with a median age of 22, three quarters of listeners 26 or younger, and just nine listeners 40 or older. Number of listens was analyzed as a categorical factor with four levels: 18–22 years old (50.5% of listeners), 23–29 years old (33.0% of listeners), 30–39 years old (7.2%), and 40+ years old (9.3%); with four levels, the analysis of listener age involved six pairwise contrasts for each of the 12 scales. Listener age yielded two credible pairwise differences, as listeners aged 40+ rated speakers across the board more *young* than did listeners aged 18–22 or 23–29. These results suggest a cumulative effect whereby listeners were rated as younger by older speakers; otherwise, listener age did not affect scale ratings.

Finally, listening device yielded just one credible difference (with 10 pairwise contrasts for each of the 12 scales), as listeners using external speakers rated speakers credibly more *young* than did listeners using headphones. As this single difference fails to suggest a real trend, however, it appears that listeners' scale ratings were not affected by listening device overall.

4.2.9. Summary: Scales

The analysis of scales responses revealed several noteworthy results. The analysis of ratings by guise indicates that Californian listeners recognize the CVS as a Californian feature. Guise also yielded almost no credible interactions with other predictors, suggesting that the perceptual effect of the CVS works independently of other speaker characteristics. Guise registered the fewest credible effects among the primary predictors, indicating that its effect on listeners' perceptions takes a back seat to the effects of other speaker characteristics. Speaker

ethnicity, on the other hand, affected all but one scale—all but *sounds like a Valley girl*, paradoxically the only scale that arguably could be construed to capture (via the original Valley girl stereotype) a raced indexicality—with Caucasian speakers generally rated more favorably. Speaker gender also substantially affected listeners' ratings.

The results by speaker region differed substantially from those of the pilot study; together with listeners' inability to reliably differentiate speakers by Californian region, this result suggests that the credible effects of speaker region in the main study were simply the result of idiosyncratic speaker characteristics rather than any features distinguishing these regions. Conversely, the relative stability of ratings by perceived region, especially claiming effects, indicates that these regions possess an evaluative reality as folk-linguistic regions. The perceived-region results indicated not only a greater set of traits differentiating Californian folk-linguistic regions from one another, but also a subset of traits that more strongly differentiate Californian folk-linguistic regions from imagined non-Californian speakers; in both sets of comparisons, speakers perceived to be from the Lower Central Valley were especially stigmatized (despite Lower Central Valley listeners representing a plurality of the listener sample), a result that mirrors the apparent erasure of the Lower Central Valley in the regional identification responses.

Put together, the effects of numerous predictors suggest interesting interrelations between the constructs represented by individual scales. For example, the "sounds like me" construct and the construct of familiarity were both subject to claiming effects and were affected by various predictors; however, listeners generally rejected the notion that speakers sounded like them (i.e., with a low baseline rating) but accepted the notion that speakers sounded familiar (i.e., with a high baseline rating). Especially complex are the interrelations between the concept of Valley

girls and other concepts. Several results indicate that among Californian listeners, the concept of Valley girls is related to but distinct from Californian-ness. Whereas both are associated to a greater degree with speakers perceived to be from California (and within California, associated with Southern California especially) as well as with the CVS, listeners generally rejected the Valley girl label and accepted the Californian label; notably, listeners also claimed speakers who sounded the most Californian, but there were no claiming effects for Valley girl ratings. Sounding like a Valley girl is also not quite the same as sounding feminine to Californian listeners, as speakers can be rated high on the former but low on the latter. Across the set of predictors, Valley girl ratings generally did not pattern with ratings of youth, relaxedness, or wealth, traits that are associated with the original popular stereotype of Valley girls. It is possible that among Californian listeners, the Valley girl stereotype has retained its association with the CVS but has lost its associations with traits that paint CVS users as ditzy, bleach-blonde beneficiaries of undue privilege; in other words, these results suggest that for Californians, the "Valley girl" stereotype may be changing to denote Californians who sound especially like California English users.

Californian listeners also assigned highly similar attributes to speakers regardless of listeners' region of origin, both in absolute terms and in terms of how listeners reacted to the CVS; the only exceptions were that Lower Central Valley listeners rated speakers overall lower for youth and higher for confidence, and Bay Area listeners attributed greater masculinity to the CVS whereas Southern California listeners attributed greater femininity to the CVS. Listeners across the state exhibited robust claiming effects for Californian-ness, familiarity, and the "sounds like me" construct. Finally, there were few notable effects of non-controlled predictors, especially given that the effect of listener mobility on ratings of youth can be attributed to the

effect of Lower Central Valley listeners. Speakers were rated less suitable for a speaking job the more that listeners listened to them and the earlier that speakers appeared in the listening task; speakers were also rated younger in earlier trials and by older listeners.

Chapter 5
Conclusion

At the beginning of this dissertation, I posed a fundamental question about social meaning in sociolinguistics: How do listeners and speakers together participate in the construction of social meaning? This dissertation has explored the social meanings of the California Vowel Shift (CVS) from the perspective of Californian listeners' reactions to variation, in an attempt to complement and complicate the claims of past sociolinguistic research that has investigated social meanings of the CVS through speakers' situational use of CVS features. These social meanings also exist against a backdrop of robust folk-linguistic stereotypes of California English that circulate beyond the state. The results of this study differ in substantial ways from the social meanings indicated by production or folk-linguistic discourses, suggesting that social meaning is the sole province of neither speakers or listeners but is constructed in the space between what is said and what is heard.

This chapter addresses these broader concerns by situating the results of this study within the context of sociolinguistic research and social meaning. I first revisit the original research

questions (5.1), I then discuss this study's broader implications (5.2), and I conclude by outlining future research directions suggested by the results of this study (5.3).

5.1. Research questions

In this section, I consider how the data in the main study bear on the original research questions:

1. To what extent (if any) do Californians perceive the California Vowel Shift as Californian?
2. What evaluations do Californians attach to the California Vowel Shift? How do these evaluations compare to popular portrayals of California speakers?
3. To what extent (if any) do listeners from different regions of California evaluate speakers of California English differently?

5.1.1. To what extent (if any) do Californians perceive the CVS as Californian?

The main study results support the claim that Californians perceive the CVS as Californian. The analysis of regional identification data indicates that non-California-shifted speakers are more likely to be identified as from outside California than California-shifted speakers. The analysis of scales data further indicates that California-shifted speakers sound more Californian to Californian listeners than non-shifted speakers; however, Californian listeners generally do not hear California-shifted speakers as sounding more like the listeners *themselves* than non-shifted speakers. In acknowledging the association between the CVS and California, Californians are unlike upper Midwesterners, who do not perceive the Northern Cities Shift as tied to Detroit (Niedzielski 1999). One possible explanation for this difference is that California English is enregistered to an extent that upper Midwestern English, Michigan English,

and Detroit English arguably are not, making it harder for Californians to hear California English as identical to conservative American English standards. I explore another possible explanation below: Californians' conception of "standardness" is simply different from that of upper Midwesterners (5.2.3. Language change). Regardless of the reason why Californian listeners hear the CVS as Californian, the finding that they do—in spite of a lack of overt folk-linguistic commentary on the CVS among Californians (Bucholtz et al. 2007)—reinforces the folk-linguistic principle that listeners may be sensitive to linguistic differences even if these differences are not overtly available to them (see 1.4.2. Perceptual dialectology, above).

5.1.2. What evaluations do Californians attach to the CVS?

The analysis of scales data with guise as a predictor indicates that, among Californians, the CVS is associated not only with sounding more Californian and more like a Valley girl, but also more confident. The positive effect of the CVS on the perception of confidence is mitigated by gender; men who use the CVS sound more confident than those who do not, whereas women sound equally confident regardless of CVS use. This result stands in contrast to discourses that attribute a lack of confidence to California speech, especially that of females—an effect often tied to the supposedly hesitant stance underlying Californian women's use of high rising terminals (Tyler 2015). Whereas high rising terminals index a lack of confidence, the CVS apparently does not.

Speakers' use of the CVS had a weaker effect on scale ratings (both in terms of the number of credible differences and the magnitude of these differences) than other speaker characteristics: gender, region (actual or perceived), and ethnicity. Guise also yielded almost no credible interaction effects with other predictors. These results suggest that the CVS does not loom large in Californians' perceptions of speakers; when Californians make judgments about

speakers' personal traits, that is, they rely primarily on characteristics such as gender and ethnicity, with speakers' use of CVS features playing only a secondary role and affecting listeners' judgments to a limited extent. Whereas other speaker characteristics interact with one another in influencing listeners' judgments, however, speakers' use of the CVS is for the most part judged independently of other speaker characteristics, a result that lends credence to the main effects of guise found in this study, as it suggests that idiosyncratic speech characteristics were not responsible for the main effects of guise.

The evaluations of the CVS that emerge from these data only partially mirror popular stereotypes of California speakers. It is somewhat surprising that Californians associate the CVS with the "Valley girl" label, as this character type—in its original conception—stands in for a negative stereotype of (Caucasian) Californians as possessing an aloof, ditzy privilege. Californians do not, however, associate the CVS with the character traits stereotypically connected to Valley girls: wealth, femininity, youth, and carefreeness—indeed, it appears that Californians do not especially associate these traits with the Valley girl construct itself. It is possible that the "Valley girl" has lost its potency as a sociolinguistic stereotype among Californian perceivers. The fact that listeners rejected the "Valley girl" label overall, however, indicates that it still retains a generally negative connotation.

The regional identification data also present evidence that Californians perceive the CVS not only as Californian, but as *Southern* Californian specifically. Whereas California-shifted speakers are equally likely as non-shifted speakers to sound like they are from the Bay Area or Lower Central Valley, California-shifted speakers are more likely than non-shifted speakers to sound like they are from Southern California; this pattern holds for listeners across the state. The perceptual link between the CVS and Southern California is not unambiguous, however, as

listeners do not hear speakers perceived to be from Southern California as more Californian than speakers perceived to be from the Bay Area, or sounding more like Valley girls than speakers perceived to be from the Lower Central Valley. Regardless, this perceptual link represents one potential ingroup social meaning of California English—California English as a Southern California phenomenon—that does echo popular portrayals of California English (such as "The Californians"). I explore this notion of California English as a Southern California phenomenon under the rubric of "centrality" below (5.2.2. Perceptual dialectology).

5.1.3. To what extent (if any) do listeners from different regions of California evaluate speakers of California English differently?

The analyses of both regional identification and scales data with listener region as a predictor reveal few effects of listener region. Listeners from the Lower Central Valley show a general preference for identifying speakers as from their home region, but listeners from across the state agree that California-shifted speakers sound more like they are from Southern California. Lower Central Valley listeners also generally hear speakers as older than do listeners from coastal regions. Coastal listeners disagree on the gendered indexicality of the CVS, with Bay Area listeners hearing the CVS as masculine and Southern California listeners hearing the CVS as feminine. Aside from gendered indexicality, the CVS has largely the same social meanings across the state.

The interaction of listener region and perceived region provides a window into how Californians construct notions of ingroup membership. Listeners from each region claim the speakers who they hear as most Californian, familiar, and sounding like them as members of their regional ingroup. In other words, listeners from the Bay Area, Lower Central Valley, and Southern California all believe that speakers from their *own* region exemplify California English

to the exclusion of other regions. This claiming effect for 'Californian-ness' is exhibited by Lower Central Valley listeners even in the face of the general stigmatization and erasure of their home region. In addition, the claiming effect for the "sounds like me" construct appears to be of a different nature than the claiming effects for Californian-ness and familiarity; whereas listeners generally did not hear speakers as sounding like them, listeners generally heard speakers as sounding Californian and familiar. Finally, the claiming results suggest an additional interregional difference: Bay Area listeners claim speakers who they perceive to be fast-speaking and friendly, whereas listeners from the Lower Central Valley and Southern California apparently only attend to Californian-ness and self-comparison traits.

5.2. Implications of this study

The results of this study shed light on previous research in ways that speak to several broader concerns in sociolinguistic theory. In this section, I consider this study's theoretical implications for several areas of sociolinguistics: social meaning (5.2.1), perceptual dialectology (5.2.2), and language change (5.2.3). I also consider the methodological contributions of this study (5.2.4).

5.2.1. Social meaning

As discussed in the Introduction chapter (1.6.1. California indexicality), previous claims about the social meanings of the CVS have been based primarily on production studies in local communities within California. These studies have claimed that the CVS carries social meanings of carefreeness (Podesva 2011), an opposition to gang affiliation (Fought 1999), and Whiteness, femininity, and privilege (Eckert 2008b). My discussion of these studies expressed uncertainty as to whether listeners pick up on and recognize the social meanings that speakers apparently utilize the CVS to transmit; in other words, do listeners' perceptions of the CVS actually indicate the

same social meanings that are suggested by speakers' production of the CVS? While the results of the present study show that listeners associate the CVS with Californian-ness, sounding like a Valley girl, and (for males) confidence, these results do not support an association between the CVS and relaxedness, wealth, or femininity, potentially casting doubt on the claims of previous production research on the social meanings of the CVS.

These two perspectives on the social meaning of the CVS can be reconciled by turning to the notion of the indexical field, a "constellation of ideologically related [social] meanings, any one of which can be activated in the situated use of the variable" (Eckert 2008a:454). The current study measured listeners' decontextualized reactions to the CVS, suggesting several core social meanings of the CVS—Californian-ness, sounding like a Valley girl, and confidence—that exist even in the absence of more specific contexts. These meanings sit at the center of the CVS's indexical field; framed in terms of Silverstein's (2003) notion of orders of indexicality, the CVS is a first-order index of Californian-ness, sounding like a Valley girl, and confidence. These core social meanings are thus potentially available for recruitment by individual speakers to index additional social meanings in narrower contexts and in more local communities.[1] As a result, the social meanings suggested by production studies of the CVS are best interpreted as indexes at the second order, third order, or beyond, building ultimately on the CVS's first-order core social meanings.[2]

[1] This interpretation lends credence to the idea that apparent homogeneity in vowel production across California masks heterogeneity in the social meanings of the CVS: "Each city has a different perceived character, and the emergence of this character as distinctive enables the speech styles of the city to become indexes of it" (Podesva et al. 2015:180).

[2] Whether listeners in these narrower contexts actually *apprehend* the social meanings intended by the speaker in using the CVS—and whether popular portrayals of California speakers are so influential as to provide, for example, the ideological links between California and relaxedness that allow the CVS to indirectly index a social meaning of carefreeness (as Podesva 2011 claims)—remain open questions.

In this way, my account of the indexical field diverges from that of Eckert's, who argues against "the traditional view of a variable as having a fixed [social] meaning…[in which] a variable is taken to 'mean' the same regardless of the context in which it is used" (2008a:464). The present research shows that, even in a decontextualized setting stripped of local understanding, the CVS does 'mean' Californian-ness, sounding like a Valley girl, and confidence to Californians—this finding strongly suggests that whatever other social meanings are attached to the CVS in perception or production are the result of ideological moves that serve to extend the indexical field outward from these core meanings. Part of Eckert's objection to what she calls "the traditional view" is that the social meanings of forms are static, almost as if imposed by fiat; I do not want core meanings to be misconstrued as static, as it seems clear that forms can lose particular social meanings, core or not. Nevertheless, it is clear from the present study that the CVS *is* associated with certain core meanings that exist outside of specific contexts.[3]

In sum, I argue based on the current research that social meaning does not reside in production alone, as speakers do not have the last word on the social meanings of linguistic forms. This research demonstrates that the indexical field, centered on core meanings that exist even in the absence of more specific contexts, crucially links production and perception in creating social meaning.

[3] Eckert's (2000) research in suburban Detroit high schools finds that speakers constructed social meanings of jock vs. burnout opposition out of the urban-associated Northern Cities Shift, suggesting that "Detroit" may have been a core social meaning of the Northern Cities Shift (see 1.5.4. Linguistic sense of place, above). If so, it is worth noting that in that study, as well as in California, macro-social place-based meanings sat at the core of these changes' indexical field.

5.2.2. Perceptual dialectology

In the Introduction chapter, I framed perceptual dialectology as the study of speakers' mental maps of dialect variation over geographical space (1.4.2. Perceptual dialectology). This research shows that the San Francisco Bay Area, Lower Central Valley, and Southern California are separate regions in Californians' mental maps of California. This finding stands in contrast to the fact that dialectological research in California has uncovered few interregional differences (aside from lexical differences) that would mark these regions as distinctive (see 1.3.1. Intrastate variation in California English, above). I argue that this perceptual state of affairs is an outgrowth of the longstanding folk notion of California human geography as divided between "NorCal" and "SoCal," which feeds into the popular contestation of the "real" California (see 1.2. Human geography of California, above). This contestation plays out in perceptions of linguistic difference, as listeners from each region believe that their region's speech uniquely exemplifies California English.[4] This finding represents an application of the language-ideological process of fractal recursivity (Irvine & Gal 2000) onto perceivers' mental maps of variation; just as human-geographic differences between the US North and South are projected onto language, creating a folk belief that Northern and Southern US dialects are different (see 1.5.1. Language ideologies, above), human-geographic differences between regions of California create a folk belief in dialect differences within the state. Notably, this process applies both in situations where there are observable differences in production (the US North vs. South) and

[4] The "NorCal" vs. "SoCal" distinction also erases inland California, which mirrors the folk-linguistic erasure found by Bucholtz et al.'s (2007) study of Californians' perceptual dialectology maps of California (see 1.6.2. California perceptual dialectology, above).

where there are not (regions of California). In sum, fractal recursivity can create perceptual-dialectological regions that have an *evaluative reality* regardless of linguistic reality.[5]

Past research has also demonstrated the importance of ingroup vs. outgroup perceptions as a theoretical construct in perceptual dialectology (see 1.4.2. Perceptual dialectology, above).This research reinforces the usefulness of investigating ingroup vs. outgroup perceptions, as the perceptions of California English circulated by popular portrayals were not reflected in Californians' evaluations of the CVS. Perhaps equally useful as a concept, but far less overtly discussed in perceptual dialectology research,[6] is what I am labeling *centrality*: the identification of speakers who are believed to most exemplify the speech of a given region. This research presents a conflicted picture of centrality in California English; on the one hand, listeners from each region of the state claimed the speakers who they heard as the most Californian-sounding, but on the other hand, listeners from across the state associated California-shifted speakers with Southern California.

The apparent contradiction between claiming and hearing the CVS as Southern Californian can be resolved via Preston's (1996a) concept of modes of folk linguistic awareness (see 1.4. Folk linguistics, above), in particular availability (the degree to which non-linguists comment upon topics of language) and detail (the specificity of folk beliefs). The CVS tends not to be the subject of overt commentary by Californians (as Bucholtz et al. 2007 shows), but

[5] This research also demonstrates that *claiming* may operate irrespective of the presence or absence of observable dialect differences. Arguably, this fact is a natural extension of the existence of perceptual-dialectological regions that do not correspond to observable dialect differences; if listeners believe that speakers from their home region are somehow different from speakers from other, "linguistically different" regions, they are liable to ascribe positive attributes to their supposed home region speakers.

[6] Among the diverse precedents for the concept of centrality are Preston's (1996b) identification of the perceived "core" of the South by reference to patterns of agreement on hand-drawn maps, Williams et al.'s (1999) asking listeners to rate speakers' perceived "Welshness," and Campbell-Kibler and Torelli's (2013) investigation of contested Ohio enregisterment on Twitter.

popular portrayals of the CVS such as Zappa's "Valley Girl" and Saturday Night Live's "The Californians" depict the CVS as a Southern California phenomenon (see 1.6.3. California enregisterment, above). The CVS thus registers low on availability to individual Californian perceivers—but in the uncommon situation in which the CVS *is* available for folk-linguistic comment, it is linked with Southern California. As a result, when operating at a lower level of folk-linguistic detail, Californians from each region of the state are free to think of their own region as central to California English; conversely, when operating at a higher level of folk-linguistic detail, Californians tend to think of Southern California as central to the CVS. Centrality may thus be mediated by the availability of certain features in folk-linguistic awareness, and this mediation can be influenced by popular portrayals that circulate among the outgroup (see 1.5.3. Enregisterment, above). To revisit a question from the Introduction, this research shows that centrality is a means by which discourses about variation may affect reactions to variation.

5.2.3. Language change

In describing the setting of this research, I argued that the CVS, as a change from below, represented a setting for probing the relationship between social meaning and language change. That some relationship should exist between social meaning and language change is intuitively clear, as speakers have an interest in using features that are associated with desirable identities, individuals, and places, and in avoiding features that suggest negative attributes. This research indicates that speech communities can indeed assign social meanings to variation of which they are unaware (or more precisely, of which their awareness is limited to lower levels of availability). Interestingly, perceivers' social valuation of a change from below does not necessarily mean that perceivers associate themselves with the change; in the ears of Californian

listeners, the CVS registers as something that 'someone else' uses. Listeners believe that the CVS exists in their local communities, and it is definitely Californian, but it is not the way that they (the listeners) speak. It is possible that this attitudinal stance is a key ingredient in allowing changes from below such as the CVS to flourish; listeners and speakers are aware of the change in the community (at least at a tacit level) but do not believe that they are participating in the change.

Often enmeshed in the relationship between social meaning and language change is the standard language ideology (see 1.5.1. Language ideologies, above); New York City English and Southern US English, for example, gradually moved from being nonrhotic to rhotic dialects in the second half of the 20th century in response to a supraregional rhotic language standard, which assigned stigmatized social meanings to nonrhoticity (Labov et al. 2006b). Some recent research has suggested that the CVS is emerging as a language standard, at least within the West (Becker et al. in press). Among Californian listeners, however, the CVS does not carry the evaluations that are the hallmark of standard English: a greater suitability for public speaking, an association with higher social class, etc. Indeed, as Fought's (2002) study of Californians' hand-drawn maps of the US suggests, it is possible that Californians differentiate "good English" (which is found in the West, especially California) and "proper English" (which is found in New England). In other words, it appears that as Californians (and perhaps Westerners in general) realign their vowel system, they are also realigning a notion of standardness away from the status traits that tend to define standardness further East.[7]

[7] This sort of ideological change also has implications for how speakers and listeners construct social meaning, which may in turn change the ways that speakers orient toward or away from incipient variation.

5.2.4. Methodological contributions

Although the use of acoustic manipulation in perceptual research is not new (D'Onofrio 2015; Fridland et al. 2004; Plichta & Preston 2005), several components of this study's methods are, to my knowledge, unprecedented in perceptual sociolinguistic research. First, this research used acoustically manipulated vowels in carrier phrases rather than single words or vowel continua. Second, the raw material for these manipulations was spontaneously produced speech, which inherently presents greater challenges for resynthesizing vowels (due to factors like the greater likelihood of nasal coarticulation), rather than careful speech. Third, vowels were manipulated for multiple speakers (both female and male) with different vowel spaces, which necessitated the creation of a procedure to calculate manipulation targets that resulted in stimuli satisfying naturalness and generalizability.

My experience attempting to create these stimuli made me realize why these novel methods had never been used before (in combination, let alone in isolation): simply put, they are very difficult to get right. Given the substantial amount of production research on vocalic variation (Labov et al. 2006b), however, it is useful for sociolinguists to have a methodological toolkit for measuring listeners' reactions to vowels in naturalistic contexts. The procedures I have outlined for determining manipulation targets (see 3.4.2. Manipulation target calculation, above) and implementing the manipulation process (see 3.4.3. Implementation of manipulation, above) expand this toolkit, as they represent improvements in the validity of matched-guise task design for studies of listeners' responses to vocalic variation. Moreover, the extension of this process to multiple speakers allows for these tasks to investigate how listeners' perceptions of vocalic variation are mediated by speaker characteristics, a crucial component in understanding how indexical fields grow outward from core meanings (5.2.1. Social meaning).

5.3. Future directions

While this study makes several contributions to sociolinguistic theory and methods, it suggests many more questions. For example, if the CVS exists in other states in the West (Oregon: Becker et al. in press; Nevada: Fridland & Kendall forthcoming; Colorado: Holland 2016), what does it mean that the CVS is available in California as a resource to index Californian identity? Going forward, I plan to continue pursuing these sorts of questions both in the near term, using data collected in this research but not analyzed for this dissertation, and beyond, through research on newly collected data.

Thus far, the analysis of production data from the 30 sociolinguistic interviews that I conducted in the interviewing stage has been limited to 12 speakers' production of TRAP and GOOSE in the retell task. I intend to conduct a broader analysis of the production data by analyzing a wider set of CVS features across all interview sections (conversation, retell, reading passage, word list, minimal pairs) and among all 30 speakers. This analysis will allow me to address hypotheses raised by the current study's perception results. For example, the finding that Bay Area listeners evaluated the CVS as masculine whereas Southern California listeners evaluated the CVS as feminine should hypothetically accompany interregional differences in speakers' use of CVS features to style gender identity (i.e., a third-wave perspective on sociolinguistic style). The fact that speakers were recorded in multiple tasks that were designed to draw differential attention to their speech will also allow me to apply a first-wave perspective on style to the production data, which touches on the question of whether Californians orient to the CVS as an emerging standard. Furthermore, this multi-part analysis will facilitate a comparison of these two sociolinguistic perspectives on style.

Beyond production, I intend to perform a separate analysis of existing discourse data. This body of data will be drawn from the interviewing stage (i.e., speakers' responses to folk-linguistic interview questions such as "Do people actually talk like Valley girls or is that all just made up?") and the main study (i.e., the open-ended responses to the question "What gave you that impression? (Be as specific about particular speech features as possible)," which followed both the regional identification and scales items). Together with the production research, this discourse study will round out the present study's findings by providing complementary perspectives on the social meanings of CVS features. Moreover, main study listeners' responses to the open-ended "What gave you that impression?" question will lend a degree of ethnographic depth to the survey's quantitative results. One listener's response of 75 out of 100 on *feminine* may be quantitatively the same as another's response of 75 on *feminine*, but not qualitatively the same, because every listener brings forth different understandings of how to listen for markers of femininity, what it means to be able to rank speakers as more vs. less feminine, etc. In other words, this discourse analysis will shed light on the ideological processes that motivate and structure the attitudinal trends uncovered by the present study.

The implications of this study also suggest several new research directions, both in language change and in folk linguistics, beyond the near term. One potential research direction arises from anecdotal observations (from audiences at presentations of my dissertation results) that hearing California-shifted versions of TRAP and GOOSE created the percept of *other* (non-manipulated) CVS features shifting in the Californian guise relative to their production in conservative guise. This observation raises the question of whether highly co-occurring features are perceived as individual features or as a single feature; that is, do listeners hear CVS vowels as a set of individual changes or all together? Future research could pursue this question via a

vowel-matching task wherein listeners hear a stimulus with shifted or non-shifted tokens of noncritical vowels (e.g., TRAP) and match a critical, non-manipulated vowel (e.g., DRESS) to one of several templates differing by the degree of shift.

Another potential direction arises from this study's finding that social meaning is sensitive to fine spectral detail, as this finding raises questions about the nature and limits of this sensitivity. One such question is how small of a change in formant frequency triggers a difference in listeners' evaluations of a speaker; that is, is there a "just evaluatable difference" that can be measured separately from a psychoacoustic just noticeable difference? Another question centers on the mathematical nature of the relationship between changes in formant frequency and changes in speaker evaluations (i.e., whether differences in speaker evaluations increase linearly, logarithmically, etc., as differences in formant frequency increase). The acoustic manipulation methods used to create the present study's stimuli could easily be extended to facilitate perceptual studies investigating these sorts of questions.

The present study's finding that Californian listeners did not believe that speakers sounded like them overall—regardless of guise—raises the question of what Californian speakers hear when they hear *themselves*. In other words, do CVS users perceive their own use of the CVS? This question could be investigated with at least two different tasks. The first task would involve Californian speakers recording a set of sentences containing tokens of a critical vowel; these tokens would then be resynthesized to several different steps on a continuum, and speakers would later attempt to identify which step is closest to their actual production. The second task would involve a shadowing task in which Californian speakers would repeat tokens of a word (which would shift formant frequency by varying amounts) as they hear them, which would reveal speakers' sensitivity to small changes in formant frequency. Of course, these types

of tasks would not have to be limited to the CVS but are suitable to investigating the status of any change in progress.

Finally, broader sociolinguistic questions about social meaning, language change, and the relationship of attitudes and identities could be addressed more comprehensively by applying the present study's general approach to social meaning to the context of Outer and Expanding Circle varieties of English. The interaction of local ideologies, media discourses, economic pressures, institutional policies, and speaker identities means that the social meanings surrounding Outer and Expanding Circle varieties are growing increasingly complex. A future study could combine perception and production research with speakers and listeners of Philippine English (PE), focusing on features such as interdental fricative stopping—which is categorically present among basilectal PE speakers, variably present among mesolectal PE speakers, and categorically absent among acrolectal PE speakers (Tayao 2004)—to investigate how the salience of these features interact both with the establishment of PE as an enregistered folk-linguistic construct within the Philippines as well as with Philippine identities as English speakers.

The findings of the present study open rich veins of potential future research in variation, perceptual dialectology, language attitudes and ideologies, and language change. This study demonstrates that regardless of the specific form of future research on social meaning, it is incumbent upon researchers to acknowledge the roles of both speaker and listener in the construction of social meaning.

References

Agha, Asif. 2003. The social life of cultural value. *Language & Communication* 23.231-73.

Aiello, Angela. 2009. A phonetic examination of California: University of California, Santa Cruz. MA thesis. Online: http://people.ucsc.edu/~gmcguir1/LabReport/AielloLRC.pdf

Allbritten, Rachael. 2011. Sounding Southern: Phonetic features and dialect perceptions. Washington, DC: Georgetown University. PhD dissertation.

Anderson, Benedict. 1983. *Imagined communities*. New York: Verso.

Associated Press. 1982, September 2. Valley girls: No way rocker's daughter talks like the record. *Palm Beach Post*, B12.

Audacity Team. 2012. Audacity [audio editing software], vers. 2.0.1. Online: http://audacity.sourceforge.net/

Baker, Wendy, and David Bowie. 2010. Religious affiliation as a correlate of linguistic behavior. *University of Pennsylvania Working Papers in Linguistics* 15.1-10.

Barras, Claude. 2002. Transcriber: A tool for segmenting, labeling and transcribing speech [freeware transcription program], vers. 1.5.1. Online: http://trans.sourceforge.net/en/presentation.php

Beal, Joan C. 2009. Enregisterment, commodification, and historical context: "Geordie" versus "sheffieldish". *American Speech* 84.138-56.

Becker, Kara. 2009. /r/ and the construction of place identity on New York City's Lower East Side. *Journal of Sociolinguistics* 13.634-58.

Becker, Kara, Anna Aden, Katelyn Best, and Haley Jacobson. in press. Variation in West Coast English: The Case of Oregon. *Speech in the western states*, ed. by Valerie Fridland, Tyler Kendall, Betsy Evans, and Alicia Wassink. Durham: Duke University Press.

Boersma, Paul, and David Weenink. 2015. Praat [phonetic analysis software], vers. 5.4.01. Online: http://www.fon.hum.uva.nl/praat/

Bourdieu, Pierre. 1977. The economics of linguistic exchanges. *Social Science Information* 16.645-68.

Bowie, David. 2013. The effect of linguistics instruction on undergraduates' linguistic attitudes. Paper presented at the Annual Meeting of the Linguistic Society of America, Boston.

Bucholtz, Mary, Nancy Bermudez, Victor Fung, Lisa Edwards, and Rosalva Vargas. 2007. Hella Nor Cal or totally So Cal? The perceptual dialectology of California. *Journal of English Linguistics* 35.325-52.

Bucholtz, Mary, Nancy Bermudez, Victor Fung, Rosalva Vargas, and Lisa Edwards. 2008. The normative North and the stigmatized South: Ideology and methodology in the perceptual dialectology of California. *Journal of English Linguistics* 36.62-87.

Bucholtz, Mary, and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7.585-614.

Campbell-Kibler, Kathryn. 2006. Listener perceptions of sociolinguistic variables: The case of (ING). Stanford, CA: Stanford University. Ph.D. dissertation.

—. 2008. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society* 37.637-59.

—. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21.135-56.

Campbell-Kibler, Kathryn, and Amber Torelli. 2013. "Bitch, I'm from Cleveland, you have the accent": Tracking enregisterment on Twitter. Paper presented at the Annual Meeting of the American Dialect Society, Boston.

Cardoso, Amanda, Lauren Hall-Lew, Yova Kementchedjhieva, and Ruaridh Purse. in press. Between California and the Pacific Northwest: The front lax vowels in San Francisco English. *Speech in the West*, vol. 1, ed. by Valerie Fridland, Betsy Evans, Tyler Kendall, and Alicia Wassink. Durham, NC: Duke University Press.

Cargile, Aaron Castelan. 1997. Attitudes toward Chinese-Accented Speech. *Journal of Language and Social Psychology* 16.434-43.

Clopper, Cynthia G., and David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32.111-40.

—. 2006. The Nationwide Speech Project: A new corpus of American English dialects. *Speech Communication* 48.633-44.

Coupland, Nikolas. 1984. Accommodation at work: some phonological data and their implications. *International Journal of the Sociology of Language* 46.49-70.

Coupland, Nikolas, Angie Williams, and Peter Garrett. 1999. "Welshness" and "Englishness" as attitudinal dimensions of English language varieties in Wales. *Handbook of perceptual dialectology*, vol. 1, ed. by Dennis R. Preston, 333-44. Amsterdam: John Benjamins.

Cresswell, Tim. 2004. *Place: A short introduction*. Malden, MA: Blackwell.

D'Onofrio, Annette. 2014. Sociolinguistic knowledge of a sound change in progress: Perceptions of California TRAP-backing. Paper presented at the Annual Meeting of the Linguistic Society of America, Minneapolis.

—. 2015. Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics* 19.241-56.

D'Onofrio, Annette, Penelope Eckert, Robert J. Podesva, Teresa Pratt, and Janneke Van Hofwegen. in press. The low vowels in California's Central Valley. *Speech in the West*, vol. 1, ed. by Valerie Fridland, Betsy Evans, Tyler Kendall, and Alicia Wassink. Durham, NC: Duke University Press.

Dailey-O'Cain, Jennifer. 2000. The sociolinguistic distribution of and attitudes toward focuser like and quotative like. *Journal of Sociolinguistics* 4.60-80.

Denwood, M. J. 2015. runjags [R package]. Online: http://runjags.sourceforge.net

Eckert, Penelope. 2000. *Linguistic variation as social practice*. Malden, MA: Blackwell.

—. 2005. Variation, convention, and social meaning. Paper presented at the Annual Meeting of the Linguistic Society of America, Oakland, CA.

—. 2008a. Variation and the indexical field. *Journal of Sociolinguistics* 12.453-76.

—. 2008b. Where do ethnolects stop? *International Journal of Bilingualism* 12.25-42.

Eckert, Penelope, and Sally McConnell-Ginet. 1992. Think practically and act locally: Language and gender as community-based practice. *Annual Review of Anthropology* 21.461-90.

Edwards, John. 1999. Refining our understanding of language attitudes. *Journal of Language and Social Psychology* 18.101-10.

Fought, Carmen. 1999. A majority sound change in a minority community: /u/-fronting in Chicano English. *Journal of Sociolinguistics* 3.5-23.

—. 2002. California students' perceptions of, you know, regions and dialects? *Handbook of perceptual dialectology*, vol. 2, ed. by Daniel Long and Dennis R. Preston, 113-34. Amsterdam: John Benjamins.

Fridland, Valerie, and Kathryn Bartlett. 2006. Correctness, pleasantness, and degree of difference ratings across regions. *American Speech* 81.358-86.

Fridland, Valerie, Kathryn Bartlett, and Roger Kreuz. 2004. Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change* 16.1-16.

Fridland, Valerie, and Tyler Kendall. forthcoming. Speech in the Silver State. *Speech in the West*, vol. 2, ed. by Valerie Fridland, Betsy Evans, Tyler Kendall, and Alicia Wassink. Durham, NC: Duke University Press.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian data analysis*, 3rd edition. Boca Raton: CRC Press.

Gelman, Andrew, and Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models*, 1st edition. Cambridge: Cambridge University Press.

Geyer, Grant. 2001. "The" freeway in Southern California. *American Speech* 76.221-24.

Giles, Howard, and Andrew C. Billings. 2004. Assessing language attitudes: Speaker evaluation studies. *The handbook of applied linguistics*, ed. by Alan Davies and Catherine Elder, 187-209. Malden, MA: Blackwell.

Giles, Howard, Richard Bourhis, Alan Lewis, and Peter Trudgill. 1974. The imposed norm hypothesis: A validation. *Quarterly Journal of Speech* 60.405-10.

Gould, Peter, and Rodney White. 1986. *Mental maps*. Boston: Allen & Unwin.

Graff, David, William Labov, and Wendell A. Harris. 1986. Testing listeners' reactions to phonological markers of ethnic identity: A new method for sociolinguistic research. *Diversity and diachrony*, ed. by David Sankoff, 45-58. Amsterdam: John Benjamins.

Gray, Tom, and Robert Scardamalia. 2012. The great California exodus: A closer look. Manhattan Institute for Policy Research. Online: http://www.manhattan-institute.org/html/cr_71.htm.

Guenter, Josh. 2000. What is English /l/ really? *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect*.113-20.

Hagiwara, Robert. 1997. Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America* 102.655-58.

Hall-Lew, Lauren. 2009. Ethnicity and phonetic variation in a San Francisco neighborhood. Stanford, CA: Stanford University. PhD dissertation.

Hartley, Laura C. 1999. A view from the West: Perceptions of U.S. dialects by Oregon residents. *Handbook of perceptual dialectology*, vol. 1, ed. by Dennis R. Preston, 315-32. Amsterdam: John Benjamins.

Hinton, Leanne, Birch Moonwomon, Sue Bremner, Herb Luthin, Mary Van Clay, Jean Lerner, and Hazel Corcoran. 1987. It's not just the Valley Girls: A study of California English. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*.117-28.

Holland, Cory. 2016. Beyond the Front Range: The Coloradan vowel space. Paper presented at the American Dialect Society, Washington, DC.

Holland, Cory Lin. 2014. Shifting or shifted? The state of California vowels. Davis, CA: University of California, Davis. PhD dissertation. Online: https://sites.google.com/site/californiapronunciation/home/findings

Honorof, Douglas N., Jill McCullough, and Barbara Somerville. 2000. Comma gets a cure. Online: http://research.ncl.ac.uk/necte2/documents/comma.pdf.

Irvine, Judith T., and Susan Gal. 2000. Language ideology and linguistic differentiation. *Regimes of language: Ideologies, polities, and identities*, ed. by Paul V. Kroskrity, 35-84. Santa Fe: School of American Research Press.

Johnson, Keith. 2003. *Acoustic & auditory phonetics*, 2nd edition. Malden, MA: Blackwell.

Johnstone, Barbara. 2009. Pittsburghese shirts: Commodification and the enregisterment of an urban dialect. *American Speech* 84.157-75.

—. 2013. "100% Authentic Pittsburgh": Sociolinguistic authenticity and the linguistics of particularity. *Indexing authenticity*, ed. by Thiemo Breyer, Véronique Lacoste, and Jakob Leimgruber. Berlin: Mouton de Gruyter.

Johnstone, Barbara, Jennifer Andrus, and Andrew E. Danielson. 2006. Mobility, indexicality, and the enregisterment of "Pittsburghese". *Journal of English Linguistics* 34.77-104.

Joos, Martin. 1948. Acoustic phonetics. *Language* 24.5-136.

Kendall, Tyler, and Erik R. Thomas. 2010. vowels [R package], vers. 1.2

Kennedy, Robert, and James Grama. 2012. Chain shifting and centralization in California vowels: An acoustic analysis. *American Speech* 87.39-56.

Kent, Raymond D., and Charles Read. 2002. *The acoustic analysis of speech*. San Diego: Singular/Thomson Learning.

Kerswill, Paul. 2002. Koineization and accommodation. *The handbook of language variation and change*, ed. by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 669-702. Oxford: Blackwell.

Kewley-Port, Diane, and Charles S. Watson. 1994. Formant-frequency discrimination for isolated English vowels. *Journal of the Acoustical Society of America* 95.

Kroch, Anthony S. 1978. Toward a theory of social dialect variation. *Language in Society* 7.17-36.

Kruschke, John K. 2015. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*, 2nd edition. London: Academic Press.

Labov, William. 1972a. Academic ignorance and black intelligence. In *The Atlantic Monthly*. Online: http://www.theatlantic.com/past/docs/issues/95sep/ets/labo.htm.

—. 1972b. On the mechanism of linguistic change. *Directions in sociolinguistics*, ed. by John J. Gumperz and Dell Hymes, 512-38. New York: Holt, Rinehart and Winston.

—. 1972c. The reflection of social processes in linguistic structures. *Sociolinguistic patterns*, 110-21. Philadelphia: University of Pennsylvania Press.

—. 1972d. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

—. 1984. Field methods of the project on linguistic change and variation. *Language in use: Readings in sociolinguistics*, ed. by John Baugh and Joel Sherzer, 28-53. Englewood Cliffs, NJ: Prentice-Hall.

—. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2.205-54.

—. 2006. *The social stratification of English in New York City*, 2nd edition. Cambridge: Cambridge University Press.

—. 2010. *Principles of linguistic change Vol. III: Cognitive and cultural factors*. Hoboken, NJ: Wiley-Blackwell.

Labov, William, Sharon Ash, Maciej Baranowski, Naomi Nagy, Tracey Weldon, and Maya Ravindranath. 2006a. Listeners' sensitivity to the frequency of sociolinguistic variables. *Penn Working Papers in Linguistics* 12.105-29.

Labov, William, Sharon Ash, and Charles Boberg. 2006b. *The Atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.
Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski, and Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15.431-63.
Ladefoged, Peter. 1996. *Elements of acoustic phonetics*, 2nd edition. Chicago: University of Chicago Press.
Lambert, Wallace E., R. C. Hodgson, R. C. Gardner, and S. Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60.44-51.
Lance, Donald M. 1999. Regional variation in subjective dialect divisions in the United States. *Handbook of perceptual dialectology*, vol. 1, ed. by Dennis R. Preston, 283-314. Amsterdam: John Benjamins.
Lippi-Green, Rosina. 2011. *English with an accent: Language, ideology and discrimination in the U.S.*, 2nd edition. New York: Routledge.
Loudermilk, Brandon. 2013. Cognitive mechanisms in the perception of sociolinguistic variation. Davis, CA: University of California, Davis. PhD dissertation.
Metcalf, Allan A. 1972. Directions of change in Southern California English. *Journal of English Linguistics* 6.28-33.
Milroy, James, and Lesley Milroy. 1999. Two nations divided by the same language? The standard language ideology in Britain and the United States. *Authority in language: Investigating standard English*, 150-60. New York: Routledge.
Montello, Daniel R., Alinda Friedman, and Daniel W. Phillips. 2014. Vague cognitive regions in geography and geographic information science. *International Journal of Geographical Information Science* 28.1802-20.
Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18.62-85.
—. 2002. Attitudes toward Midwestern American English. *Handbook of perceptual dialectology*, vol. 2, ed. by Daniel Long and Dennis R. Preston, 321-27. Amsterdam: John Benjamins.
Niedzielski, Nancy A., and Dennis R. Preston. 2003. *Folk linguistics*, 2nd edition. Berlin: Mouton de Gruyter.
Nolte, Carl. 1984, February 26. How to talk like a San Franciscan. *San Francisco Chronicle*, 14. Online: http://www.sfgate.com/entertainment/article/How-to-Talk-Like-a-San-Franciscan-3371767.php.
Ochs, Elinor. 1992. Indexing gender. *Rethinking context: Language as an interactive phenomenon*, ed. by Alessandro Duranti and Charles Goodwin, 335-58. Cambridge: Cambridge University Press.
Paveau, Marie-Anne. 2011. Do non-linguists practice linguistics? An anti-eliminative approach to folk theories. *AILA Review* 24.40-54.
Plichta, Bartek, and Dennis R. Preston. 2005. The /ay/s have it: The perception of /ay/ as a North-South stereotype in US English. *Acta Linguistica Hafniensia* 37.243-85.
Plummer, Martyn. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, vers. 4.0.0. Online: http://mcmc-jags.sourceforge.net/
Podesva, Robert J. 2011. The California Vowel Shift and gay identity. *American Speech* 86.32-51.
Podesva, Robert J., Jeremy Calder, Hsin-Chang Chen, Annette D'Onofrio, Isla Flores-Bayer, Seung Kyung Kim, and Janneke Van Hofwegen. 2013. The status of the California

Vowel shift in a non-coastal, non-urban community. Paper presented at the Annual Meeting of the American Dialect Society, Boston.
Podesva, Robert J., Annette D'Onofrio, Janneke Van Hofwegen, and Seung Kyung Kim. 2015. Country ideology and the California Vowel Shift. *Language Variation and Change* 27.157-86.
Pratt, Teresa, and Annette D'Onofrio. 2014. Awareness in enregisterment: Performance of the California Vowel Shift in SNL's "The Californians". Paper presented at the NWAV 43, Chicago.
Preston, Dennis R. 1989. Standard English spoken here: The geographical loci of linguistic norms. *Status and function of language and language varieties*, ed. by Ulrich Ammon, 324-54. Berlin: Walter de Gruyter.
—. 1996a. Whaddayaknow? The modes of folk linguistic awareness. *Language Awareness* 5.40-74.
—. 1996b. Where the worst English is spoken. *Focus on the USA*, ed. by Edgar W. Schneider, 297-360. Amsterdam: John Benjamins.
—. 1999. A language attitude approach to the perception of regional variety. *Handbook of perceptual dialectology*, vol. 1, ed. by Dennis R. Preston, 359-73. Amsterdam: John Benjamins.
—. 2011. Methods in (applied) folk linguistics: Getting into the minds of the folk. *AILA Review* 24.15-39.
Purnell, Thomas, William J. Idsardi, and John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18.10-30.
R Core Team. 2015. R: A language and environment for statistical computing, vers. 3.2.2. Online: https://www.R-project.org/
Rolle, Andrew. 1987. *California: A history*, 4th edition. Arlington Heights, IL: Harlan Davidson.
Rosenfelder, Ingrid, Joe Fruehwald, Keelan Evanini, and Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) program suite [forced phonetic aligner]. Online: http://fave.ling.upenn.edu/
Schilling, Natalie. 2013. *Sociolinguistic fieldwork*. New York: Cambridge University Press.
Silverstein, Michael. 1979. Language structure and linguistic ideology. *The elements: A parasession on linguistic units and levels*, ed. by Paul R. Clyne, William F. Hanks, and Carol L. Hofbauer, 193-247. Chicago: Chicago Linguistic Society.
—. 1992. The uses and utility of ideology: Some reflections. *Pragmatics* 2.311-23.
—. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23.193-229.
Styler, Will. 2015. Using Praat for linguistic research. Online: http://savethevowels.org/praat/UsingPraatforLinguisticResearchLatest.pdf.
Tayao, Ma Lourdes G. 2004. The evolving study of Philippine English phonology. *World Englishes* 23.77-90.
Titze, Ingo R. 2008. Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America* 123.2733-49.
Tyler, Joseph C. 2015. Expanding and mapping the indexical field: Rising pitch, the uptalk stereotype, and perceptual variation. *Journal of English Linguistics* 43.284-310.

Villarreal, Dan. in press. “Do I sound like a Valley Girl to you?” Perceptual dialectology and language attitudes in California. *Speech in the West*, vol. 1, ed. by Valerie Fridland, Tyler Kendall, Betsy Evans, and Alicia Wassink. Durham, NC: Duke University Press.

Warren, Ron, and Sean A. Fulop. 2014. The merged vowel of PIN and PEN in Bakersfield, CA. Paper presented at the 167th Meeting of the Acoustical Society of America, Providence, RI.

Wassink, Alicia. 2015. Dialect evolution in the Pacific Northwest: Reanalysis and conventionalization of a universal phonetic pattern. Paper presented at the Annual Meeting of the Linguistic Society of America, Portland, OR.

Watt, Dominic, Anne Fabricius, and Tyler Kendall. 2011. More on vowels: plotting and normalization. *Sociophonetics: A student's guide*, ed. by Marianna Di Paolo and Malcah Yaeger-Dror, 107-18. London: Routledge.

Wells, John C. 1982. *Accents of English*. Cambridge, UK: Cambridge University Press.

Williams, Angie, Peter Garrett, and Nikolas Coupland. 1999. Dialect recognition. *Handbook of perceptual dialectology*, vol. 1, ed. by Dennis R. Preston, 345-58. Amsterdam: John Benjamins.

Wolfram, Walt, and Natalie Schilling-Estes. 2006. *American speech*, 2nd edition. Malden, MA: Blackwell.

Woolard, Kathryn A. 1992. Language ideology: Issues and approaches. *Pragmatics* 2.235-49.

Appendices

Appendix A. Sociolinguistic interview conversation questions

The following questions were generated to guide conversation during the conversation section of my sociolinguistic interviews. Several of these questions are due in part or entirely to questions used in previous sociolinguistic research, such as 1.7 "Did you ever get blamed for something you didn't do?" (Labov 1984; Schilling 2013). As suggested by the college-life questions in section 2, the intended audience for many of these questions were college students, who I initially assumed would constitute the entirety of my speaker sample; I selected questions according to the interviewee. Although these questions are grouped thematically, they were by no means sequential, but rather represented a menu of options for stimulating conversation. This list of questions was never explicitly consulted during interviews. Questions were nevertheless written in a conversational register in order to encourage natural speech from the interviewee. Finally, while the extent to which I relied on pre-written questions varied from interview to interview, I always brought up at least several of the folk-linguistic questions in section 5, typically toward the end of the conversation section.

1. **Growing up**
    - 1.1. What was it like growing up around here? (Do you think it's a good place to raise kids?)
    - 1.2. Do you have a lot of family in the area?
    - 1.3. What sort of things did kids do around here when you were little? Did you have favorite playground games (freeze tag, red Rover)?
    - 1.4. Did you have any favorite TV shows when you were little?
    - 1.5. What did you do during the summers when you were little?

1.6. Did you get into trouble a lot when you were little? (What sort of things did you do that got you in trouble?)

1.7. Did you ever get blamed for something you didn't do?

1.8. Did you ever have a favorite teacher? How about a teacher you couldn't stand?

1.9. Who were your best friends when you were growing up?

1.10. Were there a lot of cliques in your high school? What were they?

1.11. What did you do with your (free) time outside of school?

1.12. Did you ever ditch class or ditch school? Did you ever do anything that your parents would totally flip out if they heard about it?

1.13. How did you decide you wanted to go to college? (Did anyone in your family go to college?)

2. Campus life/social networks

2.1. What classes are you taking right now? Are they interesting?

2.2. If someone in high school asked you whether they should apply to [school], what would you tell them? (Why? What are the pros and cons?)

2.3. Do you think it's better to go to community college first? (Why?)

2.4. What do you usually do when you're not in class? (Work, school activities, etc.?)

2.5. Do you tend to make a lot of new friends as you go along in life, or do you kind of stick with the same people?

2.6. The people you went to high school with, what are most of them doing now?

2.7. The people you usually hang out with, are they mostly people you know from classes, from work, from home, or whatever? Are they mostly your age? Is it like a racially mixed group? Do they mostly live around here?

3. Travel

3.1. If you had the chance, do you think you'd want to stick around here after you graduate? (Where would you want to go? What would you miss if you left?)

3.2. Are there any places you wouldn't want to live?

3.3. Have you spent much time in other parts of California? What are people from XYZ like?

3.4. What are the nicest parts of the state? What are the worst?

3.5. When you're out of town, what do people say when you tell them you're from ABC?

3.6. Do you think you'd ever move out of California?

4. Social media

4.1. Do you use social media sites like facebook a lot? (Which sites?)

4.2. When did you first get facebook?

4.3. Have you ever gotten into an argument with someone over facebook?

4.4. Are you friends with any of your older relatives on facebook? Do they ever comment on your posts?

4.5. Do you think social media sites are just a waste of time or are they actually useful?

4.6. When do you think is too early for a kid to get their first cell phone? What about their first smartphone?

4.7. Do you get annoyed when people drop words like "hashtag" or "lol" into conversation?

5. Language

5.1. What's some of the slang you use with your friends?

5.2. Do you think English is changing from one generation to the next? Have you noticed differences in how your friends talk compared to your parents' generation?

5.3. What kind of different dialects or accents are there around here?

5.4. Do people around here talk differently from other parts of the state? (Can you tell if someone's not from around here?/Do people from XYZ sound different from people from here?)

5.5. (If you had to rank it,) Where in California do you think people speak best? How about worst?

5.6. Sometimes in the media (like movies & TV) you get these characters from California and they're like Valley Girls or Surfer Dudes. Do people actually talk that way or is that all just made up? (Or is there any basis in reality? Why do you think these stereotypes exist?)

5.7. Do you think Californians have an accent? (Do people talk differently in California from other parts of the country?)

5.8. Where in the US do you think the most correct or standard English is spoken? Like if you were hiring someone to be a broadcaster on the national news and you wanted them to sound smart, where would you hire from? How about the worst English in the US?

5.9. How about you, do you think you have an accent? Has anyone ever told you you have an accent?

5.10. Do you ever find yourself changing the way you talk depending on the situation? Like with different people, or in more formal situations, or whatever? Do you ever try to do it on purpose? Or do you ever try to avoid sounding a certain way?

Appendix B. Reading passage, word list, and minimal pairs

Reading passage: "Comma gets a cure" (revised from the original Honorof et al. 2000 passage)[1]

Well, here's a story for you: Sarah Perry was a veterinary nurse who had been working daily at an old zoo in a deserted and dull part of town, so she was very happy to start a new job at a superb private practice around the bend from the light rail station. That area was much nearer to downtown and more to her liking. Even so, on her first morning, she couldn't help but feel stressed. She brewed a pot of coffee, ate scrambled eggs and a full bowl of oatmeal, checked herself in the mirror, and washed her face in a hurry. Then she put on a plain yellow dress and a fleece jacket, picked up her nurse's kit, put on her rain boots, and headed for work.

When she got there, there was a skinny woman with a goose waiting for her. The woman gave Sarah an official letter from the vet. The letter implied that the animal could be suffering from a rare form of foot and mouth disease. This was surprising, because normally you would only expect to see that illness in a dog or a goat. Sarah was sentimental, so this made her feel sad for the beautiful bird.

Before long, that itchy goose began to strut around the office like a lunatic, which made an unsanitary mess. The goose's owner, Mary Harrison, kept calling, "Comma, Comma," which Sarah thought was an odd choice for a name. The unruly bird was huge and really strong, so it would take some force to trap her, but Sarah had a different idea. First she tried gently stroking the goose's lower back with her palm, then singing a gentle tune to her. Finally, she held a full

[1] Comma Gets a Cure and derivative works may be used freely for any purpose without special permission, provided the present sentence and the following copyright notification accompany the passage in print, if reproduced in print, and in audio format in the case of a sound recording: Copyright 2000 Douglas N. Honorof, Jill McCullough & Barbara Somerville. All rights reserved.

bag of cool water to the goose's leg. Her efforts were not futile. A few seconds later, the goose began to tire, so Sarah was able to hold onto Comma, sit on a stool, and give her a relaxing bath.

Once Sarah had managed to bathe the goose, she wiped her off with a cloth and laid her on her right side. Then Sarah confirmed the vet's diagnosis. Almost immediately, she remembered an effective treatment from a vet school class that required her to measure out a lot of medicine. Sarah warned that this course of treatment might be expensive—either five or six times the cost of penicillin. I can't imagine paying so much, but Mrs. Harrison—a millionaire lawyer—thought it was a fair price for a cure.

Word list (105 items)

fatter	hawed	full	Kate	booting
hole	peen	Ruth	dead	huddle
language	dote	tool	trail	bone
necklace	heed	deed	boon	swimmer
beg	better	pin	two	hall
had	file	tiger	died	think
but	thigh	hoes	pan	boated
hail	body	thought	big	hill
sight	putter	foal	hayed	pun
bog	dad	cramp	that	though
dud	lot	tub	bug	hid
Peter	dating	riding	daughter	bagel
leg	fool	hell	pine	pen
writing	bloodhound	stood	did	heel
pain	head	bacon	city	poke
then	hull	road	who'd	bend
bad	when	sin	sing	thud
side	beagle	dude	bought	bull
glass	bag	Hal	these	bogus
tan	hater	hoodie	they	hide
boogie	hood	pawn	feet	roof

Minimal pairs (30 pairs)

stew/Stu	Don/Dawn	hole/hull
pool/pole	full/foal	bred/bread
had/hand	hull/hall	band/banned
cot/caught	bad/ban	wear/where
Aaron/Erin	weather/whether	fail/fool
pin/pen	when/win	writing/riding
pill/pole	root/Ruth	rap/wrap
poor/pore	fool/full	fail/fell
pan/pen	dude/did	sin/sing
caught/cat	root/rude	peel/pill

Appendix C. Demographic questionnaire for interviewees

Please ask Dan if anything is unclear!

1. What is your current level of education (some high school, high school graduate, some college, bachelor's degree, master's degree, doctoral/professional degree)?
2. What is your age?
3. Please list all the places you have lived. Underline the place you typically call your "hometown" (if any).
4. What is your gender identification (defined however you want)?
5. What is your ethnic/racial identification (defined however you want)?
6. What language(s) do you speak and how well? Include native language(s).
7. What is your major or majors (whether anticipated or declared)?
8. What is your mother's hometown, language background, and level of education?
9. What is your father's hometown, language background, and level of education?
10. Do you have any final questions or comments about this study?

Appendix D. Transcripts of main study excerpts

Following are transcripts of the 24 excerpts (2 excerpts apiece from 12 speakers) that served as the basis for the creation of matched guises in the main study. Disfluencies that were edited out of the excerpts for the main study are not represented in these transcripts; the disfluencies represented below could not be edited out (without sounding obviously edited) and were present in the stimuli that main study listeners heard. [br] = audible inbreath, [ns] = unidentifiable noise.

Exc1 The spider climbs over Stu and Matt the cat tries to get his attention that the spider's still alive he climbs on his head he climbs on his book and Stu sees the spider

Exc2 Matt the cat starts putting in his stuff his toys and scratching post and another cat and Stu comes back and takes the other cat off of the suitcase

Exc3 So Matt the cat puts all these toys in the suitcase and then he deposits a kitten in the suitcase attempting to have Stu take the kitten I believe with him on this trip

Exc4 You sort of realized what was happening because the kitten starts making this really annoying noise with one of the toys that it had removed from the suitcase and Matt the cat looks really aggravated

Exc5 Matt the cat I guess Iunno if he's just trying to pet Stu or y'know let him know that the spider's in his hair if he starts touching his hair too and then when uh Stu realizes the spider is there he kinda falls back

Exc6 Stu has a list things that he wants to pack and so he walks offscreen to go and and find those things and then Matt the cat is meanwhile putting all of his cat toys inside of the suitcase

Exc7 Matt the cat being unusually obedient for a cat [br] actually goes down onto the floor [ns] and eats up the spider and then Stu is Stu thinks he's safe and then he continues reading the book [br]

Exc8 Matt the cat comes back on top of the bed and starts playing around with the suitcase again [br] then Matt the cat gets some ideas and he starts packing some of his own things into the suitcase too

Exc9 Stu is about to go on a flight to like france and he was packing his bag and then Matt the cat saw and was like uh what are you doing I wanna get pet right now

Exc10 Stu was like get out of my suitcase and I have to pack and then the cat was like no you have to pet me or I'll put your underwear on my head and he put his underwear on his head and then Stu was like no you have to get out

Exc11 Stu took the suitcase had trouble carryin' it and then he left and then the cat was like oh no and then the kitten started playing and sound got to be really annoying

Exc12 Stu is packing and the cat wanted to like play in the suitcase and he's like no and so he took him out and put him on the bed and then he did it again then Stu did it again

Exc13 I think he bopped it on the head and the spider just took off and [br] kept on crawlin' over Stu and the cat er Matt were trying to get Stu's attention for some reason Stu didn't look at the same time the spider was

Exc14 So he tries putting the little kitty in there and then he takes off and uh Matt c- er Stu comes back and [br] takes the little kitty out and he grabs his suitcase thing

Exc15 Stu's just sittin' on the couch readin' a book and Matt the cat's just sittin' on the couch like on the edge and uh Stu sees the spider crawlin' around he starts freakin' out and he starts pointin' at it and he's tryin' to tell like Matt the cat to get it

Exc16 Matt the cat jumps back on and he starts playin' in the suitcase with all the clothes and then he has like an idea like I got to pack too and so he goes and gets some of his toys a bunch of his toys and then he gets his bed and he starts packin' those

Exc17 Stu is packin' he's goin' on a trip so he's just packin' his suitcase his suitcase is opened and um Matt the cat's tryin' to climb into the suitcase and disorganize everything

Exc18 Matt the cat starts putting some of his items in the suitcase like toys and um his tray where he puts his food and then a cat friend

Exc19 Stu was packing for a trip and he had a list and um he was just packing everything and then um Matt the cat decided hey let's hang out in the suitcase

Exc20 Stu came back and took out the other cat closed up the suitcase and he realized that it was really heavy and so he had trouble carrying it and then Matt thought that the cat was still in there

Exc21 Stu was afraid and he was trying to like get Matt to do something about it and so Matt the cat starts y'know doing what cats do and then he finally ends up eating it

Exc22 Matt the cat came kept climbing up on the bed and Stu kept telling him y'know get down get down so then Stu eventually walks away and then uh Matt comes back

Exc23 Stu is lulled into a sense of security at this point thinking his cat has more than adequately disposed of the spider and then as Stu begins to read again now now not scared of the spider

Exc24 Matt's startin' to pull stuff out startin' to pull socks out and everything and then and then Stu walks out of the room and when Stu walks out of the room uh Matt takes the opportunity to pack himself a nice bag

Appendix E. Implementation of acoustic manipulation

The basic procedure for vowel resynthesis, which relies upon the source–filter theory of speech production, is as follows: linear predictive coding (LPC) is performed on a vowel to estimate the acoustic effects of the filter (formants), the vowel is inverse-filtered to derive the underlying source, the filter is modified, and the source is passed through the modified filter. In order to be implemented in this study and preserve naturalness, however, this basic procedure required several adaptations, as it is relatively easy to produce obviously computer-like sounds with this process. The need for adaptations was especially the case given that the process had to account for 152 tokens (each manipulated twice), some of which were rather far from their target values, all produced in spontaneous speech by 12 speakers with differing vocal tract characteristics and embedded within carrier phrases. In order to ensure that the process was replicable and that it applied evenly to all 152 tokens, I coded several Praat scripts (available upon request) that took as input the 24 excerpts, along with data on the timing, formant measurements, and targets of each token, and produced two versions of each excerpt: one conservative guise and one Californian guise.

The acoustic manipulation proceeded according to the following steps for each token:

1. The token was extracted from the excerpt with buffers of 50 ms at both edges (later manipulation steps would trim these buffers). For the purposes of acoustic manipulation, tokens were defined as the TRAP/GOOSE vowel plus any preceding and/or following sonorant (except /r/) within the same word. Given that sonorant consonants have formant structure (Johnson 2003), these vowel-adjacent sonorants were included in the manipulation in order to avoid abrupt "jumps" between the F2 of the unmodified sonorant and the modified vowel. For some vowels, a neighboring

sonorant reacted poorly to the manipulation; in these cases, the sonorant was simply excluded from the token (i.e., they were not manipulated), with no appreciable negative effect on naturalness. Neighboring /r/s were excluded from all tokens because /r/ regularly resisted being manipulated cleanly.

2. The token was separated into two components, a high-frequency component and a low-frequency component, to facilitate the LPC calculation. Because LPC calculations find a user-specified number of formants within a given frequency range (and because the spectral information that is most informative about vowel location is relatively low in frequency), it is easier to perform an LPC over a smaller low-frequency range than attempting to count formants in higher frequencies, where formants' low amplitude makes them difficult to detect. The high-frequency component was created by high-pass filtering the original token above a maximum frequency, which was 5000 Hz for almost all tokens and 5500 Hz for others.[2] The low-frequency component was created by downsampling the original token from 44100 Hz to twice the maximum frequency, effectively low-pass filtering the original token.
3. Praat computed an LPC on the downsampled token, with a prediction order (i.e., the number of coefficients to use in the LPC calculation) of twice the number of formants. As with the maximum frequency, the number of formants had to be specified for each token in order to return the LPC that best modeled the token's formant structure. Most TRAP tokens were best modeled with 4.5 formants (i.e., 9

[2] It was necessary to specify the maximum frequency for each token; for some tokens, the LPC calculations failed (i.e., gave bad results) on the lowest 5000 Hz of the signal but succeeded on the lowest 5500 Hz, likely due to a spectral amplitude peak around 5000 Hz (i.e., a high formant).

LPC coefficients) below the maximum frequency and most GOOSE tokens were best modeled with 5 formants, but some tokens of TRAP and GOOSE used 4 or 5.5 formants. Praat's standard values were used for the other LPC input parameters: 25 ms window length, 5 ms time step, and 50 Hz pre-emphasis frequency.

4. The glottal source was derived by inverse-filtering the downsampled token through the LPC (the filter).
5. The LPC was converted to a Formant object, a type of Praat object that stores the frequency and bandwidth of each formant at each time point. (In Praat, Formant objects store the same basic data as LPC objects, but Formant objects are far easier to directly alter than LPC objects.)
6. **Main manipulation step:** The Formant object was adjusted iteratively so that the midpoint F2 of the *vowel* (i.e., excluding a neighboring sonorant, even if the sonorant was included in the token) matched the relevant target value within an acceptable margin of error.[3] This step utilized a Praat function by which a constant can be added to the value of a formant at all timesteps (e.g., "increase F2 by 70 Hz along the entire trajectory").[4] One issue with this approach is that once the source is filtered through

[3] In an earlier implementation of the acoustic manipulation process, this adjustment step was performed non-iteratively (e.g., if a given vowel had an original F2 of 1582 Hz and its target value was 1322 Hz, the function was used once to subtract 260 Hz from F2), but the result was sometimes wildly off target, especially when the adjustment was large (200 Hz or more). Hence, the final version of this manipulation process implemented the main manipulation step iteratively rather than all at once.

[4] Although Praat's function for altering Formant objects allows for more complex operations than simply adding or subtracting a constant (e.g., the linear function used in formant transition smoothing in step 13), I determined that this simple additive method would be best for the main manipulation step. One advantage of this method is that it preserves the original trajectory of the formants, which is useful because formant transitions serve as an acoustic cue to neighboring stops' place of articulation (Johnson 2003). More importantly, this method could be easily applied to all of the tokens in the excerpts, rather than attempting to come up with different manipulation formulae for a variety of cases. In addition, the tokens produced by this method satisfied the criteria of naturalness and generalizability upon auditory inspection; similarly, in synthesizing vowel continua based on actual speakers' vowels, Fridland et al. (2004:7) found that "the most natural sounding tokens resulted from the same degree of formant change (ΔF) along the entire vowel trajectory."

the adjusted Formant object, there is no guarantee that the new formants will fall exactly where they are specified to fall; for example, if the Formant object is adjusted so that F2 is increased by 70 Hz, the resultant sound may have an F2 that is 50 Hz higher, 68 Hz higher, or 84 Hz higher than the original. In short, the 'messy' nature of speech sounds (especially those produced spontaneously)—not to mention the messiness of Praat's estimates of formant frequencies—precludes a hyper-precise adjustment of formants in resynthesized vowels.[5] Fortunately, hyper-precision is not necessary because there is a lower limit to which human perceivers can detect differences in formant frequencies (or other types of sensory stimuli): the just noticeable difference (JND). Manipulated vowels were deemed acceptably close to the target value if they fell within 1 JND of the target, defined as 33.09 Hz for TRAP F2 and 21.86 Hz for GOOSE F2 (Kewley-Port & Watson 1994:492).

7. After each iteration of the main manipulation step, the difference between the current and target F2 was measured and compared to the JND for that vowel. If the difference exceeded 1 JND, the Formant object's F2 values were adjusted by 1 JND. The source was then filtered through the adjusted Formant, F2 was measured for the resultant sound, and this F2 value served as the input to the next iteration of the main manipulation step.[6] If the difference between the current and target F2 was less than 1

It should be noted, however, that because of the messiness of formant manipulation (see fn. 5), it was not always the case that formants moved by the same amount along their trajectory.

[5] This 'messiness' of the manipulation is due not to shortcomings in Praat's implementation but because it is impossible to completely decouple the filter from the source (Titze 2008). Moreover, LPCs are imperfect representations of the vocal tract filter. The calculation of LPCs assumes that there are no antiresonances (zeroes) in the signal; as a result, zeroes are not accounted for in the filter and thus remain in the source after inverse-filtering (Ladefoged 1996). As a result, if a formant is manipulated to be near the frequency of a zero that remains in the source, its amplitude will be greatly diminished.

[6] The resultant sound was used only for measuring the change in F2, as the change from one iteration to the next was never exactly equal to 1 JND (for reasons of messiness discussed in fn. 5). The alternative would have been to use this manipulated sound as the basis for the following iteration of altering F2, but an earlier implementation of this

JND, the Formant object's F2 values were adjusted by this remaining difference. At times, the actual F2 value of the sound resulting from this final iteration was further from the target than that of the penultimate iteration; in these cases, the penultimate version of the filter was accepted as the final version.

8. Once the resultant vowel was acceptably close to the target value (within 1 JND, as determined iteratively), the source was passed through the modified filter.
9. As mentioned in the section on manipulation target calculation (3.4.2), eight Californian tokens of GOOSE and one conservative token of TRAP were set to have their F3 values raised in order to avoid F2 overlapping with F3. For these nine tokens, steps 6–8 were performed on F3 prior to F2.
10. After F2 was adjusted and the source was passed through the modified filter (step 8), the new, manipulated low-frequency component was resampled back to 44100 Hz and combined with the (unchanged) high-frequency component, creating a manipulated token with a sampling frequency of 44100 Hz. (The manipulated low-frequency component had no information above the maximum frequency before or after upsampling, but its sampling frequency had to match that of the high-frequency component in order to combine the two sounds.)
11. The manipulation of the token's formants also affected the token's intensity (amplitude), in some cases causing abrupt mid-vowel rises or drops in amplitude that would be articulatorily impossible in speech. To mitigate this effect, the token was

step using this procedure was unworkable. After numerous iterations, the eventual sound was often highly degraded, as the process of calculating LPCs, inverse-filtering to derive the source, and passing the source through the modified filter produced noise that was multiplied with each iteration. As a result, the final version of this process iteratively modified only the filter itself, thus minimizing the amount of noise introduced by inverse-filtering and filtering.

subjected to a script that copied the intensity contour of the original token onto the manipulated token, then scaled the manipulated token to the mean intensity of the original.[7] This step substantially improved manipulated tokens' naturalness.

12. The manipulated token was stripped of the 50-ms buffers and spliced back into either the conservative or Californian guise.
13. Issues of discontinuous formant transitions between manipulated vowels and non-manipulated neighboring sonorants were mitigated by smoothing these transitions. In 14 cases, manipulating a GOOSE token that was adjacent to a non-manipulated sonorant (e.g., *Stu walks*) resulted in a discontinuous "jump" in F2. To eliminate these jumps, F2 was measured at the start and end points of a 40-ms clip centering on the token–sonorant boundary (i.e., either the final 20 ms of the preceding sonorant and the first 20 ms of the token, or the final 20 ms of the token and the first 20 ms of the following sonorant). F2 was then re-defined as a linear function over the 40-ms clip connecting these start and end points. For example, if F2 was 1552 Hz at the start point and 992 Hz at the end point, a change of −560 Hz over 40 ms, F2 was set to start at 1552 Hz and decrease by a slope of 14 Hz/ms. This clip was subjected to the same process outlined in steps 1–12, except that the main manipulation step (step 6) was performed just once, not iteratively. (Note that this smoothing procedure, which only affected the first or final 20 ms of the token, had no effect on how far the vowel was from the target, which was measured at the vowel midpoint.) Although this step bolstered the naturalness of stimuli in theory, in actuality I was only able to detect an increase in naturalness for smoothed over non-smoothed F2 transitions in one of the

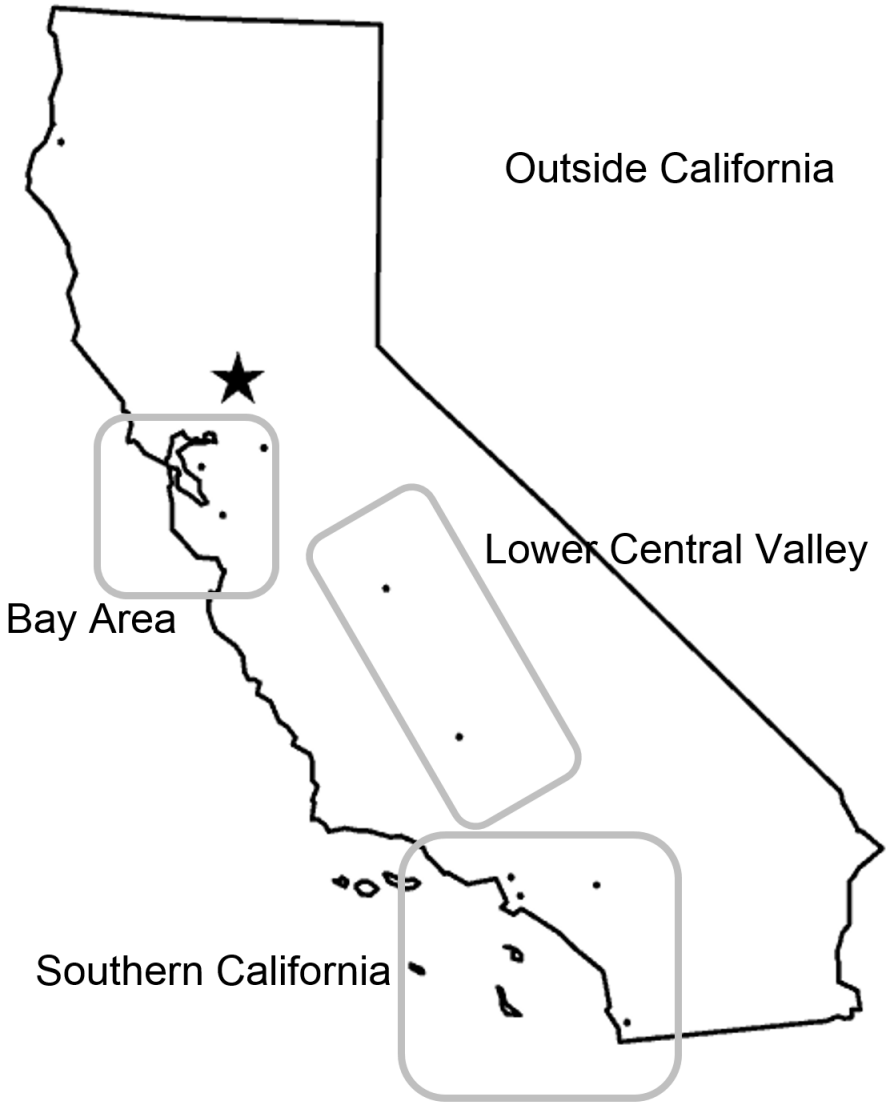
[7] Pitch was identical in the original and manipulated tokens, given that pitch is a property of the glottal source and the source was not modified.

14 cases; in the others, the F2 transitions sounded sufficiently natural regardless of smoothing.

14. Once all tokens were manipulated for a given stimulus, the average intensity of the stimulus was scaled to 65 dB so all stimuli would have the same loudness.

Prior to manipulation, five problematic tokens (one TRAP, four GOOSE) were replaced in their original excerpts. After the acoustic manipulation process was run for all 24 excerpts, a trained phonetician listened to the manipulated stimuli to gauge naturalness and generalizability. Stimuli were deemed to be satisfactory after several small adjustments.

Appendix F. Response map for main study



Appendix G. Demographic questionnaire for main study listeners

1. What is your age?
2. What is your current occupation?
3. What is your gender identification?
4. What is your ethnic/racial identification?
5. Please list all the places you have lived (for at least a month at a time, not counting vacations) and what ages you lived there.
6. What do you consider your "hometown"?
7. Where do you currently live?
8. Do you consider yourself to be a Californian, and why?
9. How well do you speak English?
10. How many language(s) do you speak other than English?
11. What many language(s) do you speak other than English?
12. What language are you most comfortable speaking?
13. Have you ever studied linguistics?
14. Do you think you speak with an accent? If so, how would you describe it?
15. Do you think there's such a thing as a "California accent"? If so, how would you describe it?
16. How did you listen to the sound clips in this survey? (Earbuds, headphones, internal computer speakers, external speakers, etc.)
17. How did you hear about this survey?
18. Do you have any final questions or comments about this survey?

19. Please enter your email address. (This will only be used for a confirmation email and to contact you if you win one of the Amazon.com gift cards. We will **not** contact you for any other reason.)
20. If you would like to receive information about the results of this study, click this check box.
21. If you are interested in being contacted for future research, click this check box. This is completely optional.
22. If you would like to help us out by uploading a clip of yourself reading a short story, click this check box then click "NEXT ≫" (this will not affect your eligibility for the gift card drawing). Otherwise, just click "NEXT ≫" to complete the survey.