Article

Journal of English Linguistics
2018, Vol. 46(1) 52–78

DOI: 10.1177/0075424217753520
journals.sagepub.com/home/eng
SAGE

The Construction of Social Meaning: A Matched-Guise Investigation of the California Vowel Shift

Dan Villarreal[1]

Abstract

Past production research on the California Vowel Shift (CVS) has suggested that the CVS carries social meanings of carefreeness, Whiteness, femininity, and privilege (e.g., Eckert 2008b), but it is unclear whether these social meanings reflect listener perceptions. In the present study, Californian listeners heard speech samples, guessed where speakers were from, and rated speakers on language attitudes scales; stimuli in this task were matched guises differing by speakers' use of two CVS features. The results indicated that listeners associate these features with Californianness, sounding like a Valley girl, and (for male speakers) confidence, complicating the social meanings suggested by production studies. I discuss these results in terms of how interaction context guides the perception of social meaning by activating subsets of the indexical field. This research also introduces two innovative methods for investigating sociolinguistic perception: stimuli created using resynthesized vowels within spontaneous speech produced by multiple speakers, and statistical inference via Bayesian hierarchical modeling.

Keywords

sociolinguistics, American dialects, language attitudes, perceptual dialectology, phonetics

1. Introduction

The study of social meaning—which connects patterns of language variation with the wider social world—has benefited from foci on speakers' use of variation in ongoing

[1]University of Canterbury, Christchurch, New Zealand

Corresponding Author:
Dan Villarreal, University of Canterbury, Christchurch, New Zealand.
Email: djvill@ucdavis.edu

interaction (Eckert 2008b; Fought 1999; Podesva 2011), perceivers' decontextualized notions of how variation maps onto geography (Bucholtz et al. 2007; Preston 2011), and listeners' perceptions of variation (Pharao et al. 2014; Tyler 2015; Watson & Clark 2013). Research of this last type has demonstrated that listeners play an active role in the construction of social meaning, underscoring the importance of perception-based research on social meaning (Campbell-Kibler 2007; Soukup 2013). A number of challenges remain for this research, including challenges of methods, such as how to generate valid stimuli, and of interpretation, such as how to reconcile social meanings apparent in production with those in perception when the two diverge. The present study focuses on listener perceptions of two features of the California Vowel Shift (CVS), which has been studied from the perspective of production but not perception (save for Villarreal 2016b), in order to address how listeners and speakers together participate in the construction of social meaning. In so doing, I address several of the challenges, both methodological and theoretical, inherent to perception research on social meaning.

The systematic study of social meaning in variationist sociolinguistics originated partially as a response to shortcomings in the correlational approach to sociolinguistic variation that characterized the "first wave" of variation study (Eckert 2012). Under this correlational approach, variation in linguistic forms is associated with speakers' static membership in macrosociological categories such as those based on ethnicity, gender, or social class. More recent "third-wave" research has approached sociolinguistic variation and social meaning through the lens of "indexicality" and in particular the multidimensional "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). Under this account, the social meanings of a given variable are not fixed but subject to ideological mediation in ways that draw upon pre-existing social meanings within the indexical field to extend the indexical field outward. For example, hyperarticulated /t/ release, in which word-final /t/ is released with a stop burst rather than glottalized, deleted, or flapped, is associated with clarity and emphasis in American English; in turn, ideological associations between clarity and qualities such as education and refinement allow speakers to employ /t/ release to index a wide array of social types in different contexts: nerd girls, Orthodox Jewish boys, gay divas, and others (Eckert 2008a). At the same time, speakers' agency in using variation to project identities is constrained and limited by the extent to which listeners pick up on these social meanings (Schilling 2013); that is, we should avoid uncritically basing claims about social meanings on speakers' stylistic use of variable forms.

Soukup (2013) provides a framework for social meanings that includes a consideration of listeners' role in the construction of social meanings, arguing that listener uptake of social meanings requires two ingredients: listeners' recognition of the contrastiveness of linguistic forms, and listeners' association of contrasting forms with contrasting social meanings. I offer two examples, one positive and one negative, to illustrate this framework. In one example, Soukup (2013) assessed a proposition emerging from production research that Austrian speakers use Austrian German dialect (*Dialekt*) features to index negative social meanings. The results of perceptual tasks revealed first that listeners associate certain features with a gestalt perception of *Dialekt*—indicating their recognition of contrasting forms—and second that listeners

evaluate recordings containing these features as less educated, intelligent, or sophisticated—indicating an association of contrasting forms with contrasting meanings. These findings indicate that the negative social meanings of *Dialekt* do not exist only in speakers' production, but also in listeners' perception. In another example, Niedzielski (1999) found that Michigan listeners were more likely to match a fronted token of *pop* with a hyper-standard vowel (further back than canonical /ɑ/) than with the actual Northern Cities Shifted token when told that the speaker was from Michigan. Despite the fact that Michigan speakers used features of the Northern Cities Shift to construct locally meaningful identities in production (Eckert 2000), Michigan listeners in Niedzielski's (1999) study failed to hear a contrast between shifted and standard /ɑ/, indicating some limit to listener uptake of the meanings present in production.

The present research continues in this vein by bringing a listener-based perspective to bear on the social meanings of the CVS, for which previous production research has suggested meanings such as carefreeness (Podesva 2011) and femininity (Eckert 2008b). In keeping with Soukup's (2013) framework, the present research addresses whether Californian listeners recognize CVS features as contrasting with more conservative American English vowels, how their perceptions of CVS features compare to the social meanings that have been suggested in previous production research, and, crucially, how any differences between the social meanings emerging from production versus perception can be reconciled.

In this study, Californian listeners heard samples of speakers from different regions of the state, attempted to identify the regional origin of the speaker, and rated the speaker on affective scales. Each stimulus belonged to one of two matched guises differing only by the speaker's use of California-shifted or non-shifted vowels. The results reveal differences between the social meanings of two CVS features indicated by listener perceptions in this context—associations with California, sounding like a "Valley girl," and confidence—versus those suggested by production research on the CVS. In the concluding section, I consider two accounts for this disconnect: either certain meanings of the CVS are present in production but not perception, or the contextualization cues (Gumperz 1982) present in the stimuli and the wider task allow for the activation of meanings like confidence but not others.

2. California and California English

California is by far the most populous state in the United States, though California's population is unevenly distributed between densely populated urban areas and wide expanses of thinly populated terrain. California's two most populous areas, together accounting for well over half of the state's population, are Greater Los Angeles in the south and the San Francisco Bay Area in the mid-north. These two 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 (Montello, Friedman & Phillips 2014). While both Greater Los Angeles and the Bay Area are coastal, the majority of non-coastal Californian population is in the Central Valley, a vast agricultural region dotted with numerous metropolitan areas.

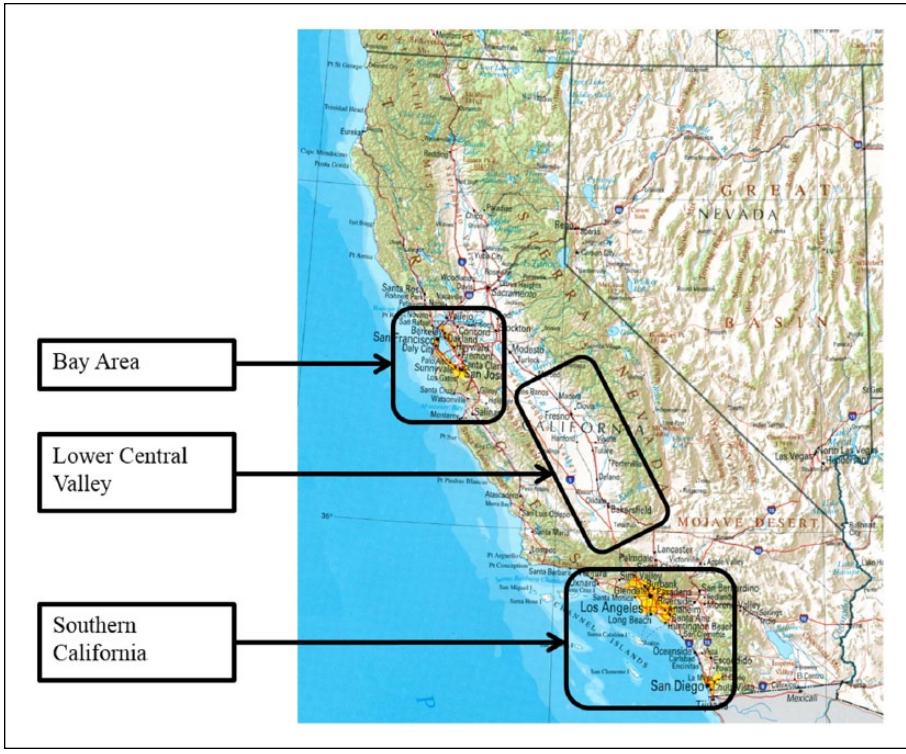

Figure 1. Californian Regions in This Study (Adapted from Public Domain Map at http://commons.wikimedia.org/wiki/File:California_ref_2001.jpg)

In this study, all stimulus speakers and most perceptual task listeners were drawn from three regions of California: the Bay Area, the Lower Central Valley, and Southern California (Figure 1); these were also the response regions in the perceptual task (along with an "outside California" option). These regions differ along economic, geographic, and demographic lines; the agriculture-based economy of the poorer, less-urbanized Lower Central Valley contrasts with the more diverse economies of the wealthier, denser coastal regions. Moreover, despite the scant evidence for intrastate differences in California English (as discussed in 2.1), these regions enjoy differential folk-linguistic status. Southern California linguistic stereotypes tend to stand in for the state as a whole (e.g., Podesva 2011), and Californians themselves project the "NorCal"/"SoCal" dichotomy onto folk-linguistic difference, often erasing the Lower Central Valley (Bucholtz et al. 2007).

2.1. English Language Variation in California

Numerous studies have uncovered a California Vowel Shift (CVS) consisting of several subsystems: the LOT/THOUGHT merger, high and mid back vowels fronting (e.g., GOOSE fronting), front lax vowels backing and/or lowering (e.g., TRAP backing, KIT lowering),

and TRAP-N (prenasal TRAP) raising (Eckert 2008b; Kennedy & Grama 2012); TRAP thus exhibits a nasal split, raising prenasally and backing elsewhere.[1] A consequence of these vowel movements is that whereas GOOSE is traditionally further back than TRAP in American English, (non-prenasal) TRAP is further back than GOOSE in California English. In actuality, TRAP, GOOSE, and many features of the CVS are prevalent throughout the Western US (Fridland, Bartlett & Kreuz 2016) and beyond (Labov, Ash & Boberg 2006). Nevertheless, media portrayals of California speakers (e.g., Frank Zappa's "Valley Girl," Saturday Night Live's "The Californians") reinforce in the popular imagination a link between CVS features and Californian ways of life, so in accordance with these attitudes I will treat the CVS as a Californian phenomenon even though its reach is broader in sociolinguistic reality. Finally, the CVS appears to be subject to little regional variation within California; past research has uncovered only minor vocalic variation between regions of California (Kennedy & Grama 2012; Warren & Fulop 2014; D'Onofrio et al. 2016), and Californian listeners are unable to reliably place the origin of Californian speakers based on their speech (Villarreal 2016b).

2.2. California Indexicality

Several studies of CVS features have suggested a variety of indexical meanings for California vowels. First, Podesva (2011) finds that "Regan," a Southern Californian man whose speech was recorded in different social settings, utilizes several CVS features to a greater extent in a party situation than in a professional meeting. Podesva (2011) argues that Regan's use of California vowels indexes a "partier" persona, which is connected to larger social meanings of Californians as carefree and fun. Second, among Chicano English speakers in Los Angeles, Fought (1999) finds that the social factor most explanatory of GOOSE fronting was not ethnicity or social class, but gang affiliation, with non-gang individuals more likely to front and gang-affiliated individuals less likely to front. Third, a study of the TRAP nasal split in Bay Area elementary schools suggests that the CVS has been "gendered, raced, and classed" to index Whiteness, femininity, carefreeness, and privilege (Eckert 2008b:29). Finally, an earlier version of the current study found no correlation between the degree to which speakers used CVS features and listeners' ratings (Villarreal 2016b); this earlier study did not directly control for speakers' use of CVS features, however, necessitating a more sensitive measure of how the social meanings of the CVS manifest in perception.

These studies leave open the question of whether listeners actually apprehend the social meanings that speakers evidently transmit in using CVS variables (Schilling 2013; Soukup 2013), as we lack perceptual research that demonstrates the existence of these indexical meanings in listeners' perceptions of the CVS.

3. Methods

The present research seeks to better understand how Californian listeners participate in the construction of the California Vowel Shift's social meanings by addressing the following research questions:

1. Do Californian listeners recognize a contrast between CVS and non-CVS vowels?
2. What social meanings do Californian listeners attach to CVS features in perception? How do these compare to the social meanings that are suggested by speakers' production of the CVS?

These research questions were investigated via a task that combined methods from perceptual dialectology—dialect recognition tasks (Williams, Garrett & Coupland 1999)—and language attitudes—matched-guise techniques (Campbell-Kibler 2007). In this task, Californian listeners heard samples of spontaneously produced speech, guessed where speakers were from, and rated speakers on affective semantic differential scales. Each speech sample represented one of two matched guises (stimuli differing only in one crucial feature): either a California-shifted guise or a conservative (non-shifted) guise. As described below, only a subset of CVS features was modified. Scripts coded in Praat (Boersma & Weenink 2015) created these acoustically manipulated guises via vowel resynthesis.

This task was modeled in part on a dialect recognition study in which Welsh listeners rated speakers from across Wales on affective scales (e.g., likeability) and guessed speakers' regional origin (Williams, Garrett & Coupland 1999). Speakers who were rated most likeable were also more likely to be identified as belonging to a listener's regional ingroup—whether or not this was actually true; for example, Cardiff listeners found Northwest speaker 2 more likeable than Northwest speaker 1 and in turn misidentified only Northwest speaker 2 as a Cardiffian (Williams, Garrett & Coupland 1999:356). These results indicate that when listeners judge where a speaker is from, they do not simply match the speaker to preexisting acoustic templates of regional speech; they also draw on their attitudes toward certain regions.

3.1. Stimuli

Forty-eight stimuli were created from the speech of twelve Californians, each of whom uniquely represented a region (Bay Area, Lower Central Valley, Southern California) × gender (female, male) × ethnicity (Latina/o, Caucasian) cell. Each speaker had lived in their respective region for their entire life, spoke English natively, and was between nineteen and thirty years old. A separate analysis of speakers' production in careful interview tasks indicated that all stimulus speakers exhibited the CVS (though speakers exhibited only moderate GOAT fronting); there were no notable differences in speakers' production of CVS vowels as a function of speakers' region, gender, and/or ethnicity. Two excerpts per speaker were drawn from spontaneously produced retellings of events presented in cartoons. Each excerpt was acoustically manipulated to create a California-shifted guise and a conservative (non-shifted) guise.

Rather than manipulate all nine features of the CVS, just two features were manipulated: TRAP backing and GOOSE fronting. Utilizing two features as a substitute for the entire shift was necessary to ensure a certain level of experimental control over the stimuli; ensuring that a sufficient number of tokens of every CVS feature

was represented in each stimulus would require eliciting much longer stimuli and/or compromising the spontaneity with which stimuli were produced. As a result, only TRAP and GOOSE varied between guises of the same stimulus, as all other vowels were kept at their original formant frequencies, and—aside from the removal of disfluencies and *like* tokens (see below)—no other features of the speakers' speech were manipulated. As previously mentioned, neither TRAP nor GOOSE is unique to the CVS (Labov, Ash & Boberg 2006). However, popular portrayals of California English such as Saturday Night Live's "The Californians" sketches have incorporated both features (Pratt & D'Onofrio 2014), so they are arguably associated with California English at some level of public awareness—though this is not to say that TRAP and GOOSE are uniquely "enregistered" as markers of California English.[2] Furthermore, it is likely that the presence of other, non-manipulated CVS features affected listeners' responses to stimuli; due to space considerations, however, an analysis of these effects will be left to future study. As a result, this current study's findings will be framed in terms of TRAP/GOOSE rather than the CVS more broadly.

Excerpts were drawn from cartoon retells that took place during sociolinguistic interviews conducted by the author, a non-Californian. In this task, speakers recounted humorous videos featuring the antics of a mischievous cat and his owner, who were given the names "Matt the cat" and "Stu" in order to seed TRAP/GOOSE tokens. This design ensured that excerpts were spontaneously produced, thus capturing a broader range of interspeaker variation than would be elicited by a read-speech task. Moreover, while it is doubtful that any stimulus content can ever be truly "neutral" (Campbell-Kibler 2009), this cartoon-retell method limited content variation from one excerpt to another without requiring the speakers to read identical texts (see Appendix A for excerpt texts).

Each excerpt contained two to six tokens apiece of TRAP and GOOSE (means: 3.8 TRAP tokens and 2.6 GOOSE tokens), pertained to the cartoon itself, and was plausible as a syntactic and prosodic stand-alone unit. Because excerpts varied in fluency, fourteen of the twenty-four excerpts were edited in Praat to remove disfluencies; three linguistically trained listeners could not distinguish the edited excerpts from the non-edited excerpts. All excerpts were pretested with untrained Californian listeners to ascertain excerpts' naturalness and relative evaluative neutrality; pretesting listeners heard unedited excerpts and listed personal characteristics of the speaker in an open-ended fashion. These listeners could not detect that some excerpts had been edited. While pretesting listeners judged excerpts to be relatively neutral overall, they did attend to speakers' use of *like*, a stereotype of California English (Dailey-O'Cain 2000), so five excerpts were further edited to remove syntactically optional *like* tokens. The final stimuli were 9.5–12 seconds long (mean: 10.7 seconds).

This study's methods for creating stimuli built on those of past sociolinguistic research in several ways. Past matched-guise research has used splicing techniques to embed variables like (ing) (Campbell-Kibler 2007) and *like* (Dailey-O'Cain 2000), but not vowels, in carrier phrases. Past perceptual research has utilized acoustically manipulated vowels in stimuli, but—with the exception of a study featuring manipulated vowels in reading passages by a single speaker (Watson & Clark 2013)—these

stimuli have been limited to vowel continua or single words produced in a reading style (e.g., Fridland, Bartlett & Kreuz 2004), not speech that listeners would interpret as occurring in discourse. The reason for this gap is simple: it is rather difficult to make acoustically manipulated vowels sound natural in spontaneous speech for multiple speakers. As this study's acoustic manipulation process represents a novel methodological tool in sociophonetics, I discuss the process in some detail in hopes of helping readers to surmount the difficulties latent in using resynthesized vowels within spontaneous speech produced by multiple speakers.

The acoustic manipulation process consisted of two stages: the determination of "manipulation targets"—the formant values representing conservative versus California-shifted TRAP and GOOSE tokens—and the vowel resynthesis itself. The challenge inherent in finding appropriate targets is twofold. Individual speakers vary considerably in the size of their vowel spaces, meaning a given feature's manipulation target must be defined relative to individual speakers' vowel spaces. Vowel space variation notwithstanding, it is still impossible to identify, a priori, a target that exemplifies a certain feature to a greater extent than any other possible target (e.g., there is no prototypically fronted GOOSE).

In order to address this challenge, this study relied on speakers' natural ranges of vowel variation, basing manipulation targets on the production of individual speakers and the overall speaker group, thus ensuring that each individual speaker's manipulated vowels not only were plausible productions for the speaker but were also roughly comparable to other speakers' manipulated vowels. Conservative (i.e., fronter) TRAP targets for each speaker were calculated as the average of the speaker's second highest TRAP F2 value (i.e., second most front TRAP token) within the retell task and the average of all same-gender speakers' second highest TRAP F2 values within the retell task. (Second highest/second lowest values were used in place of highest/lowest values to mitigate the effect of outliers.) For example, the conservative TRAP target for female speaker CV03-NG was 1943 Hz, calculated by averaging her second highest TRAP F2 value (1990 Hz) and the average of all female speakers' second highest TRAP F2 values (1895 Hz). This procedure meant that, in effect, each speaker's conservative target was a weighted average of their extreme TRAP F2 productions and those of their same-gender peers. Californian (i.e., backer) TRAP targets were then calculated by subtracting the average of same-gender speakers' ranges of TRAP production (females: 269 Hz; males: 300 Hz) from each speaker's conservative TRAP target; as a result, all speakers of the same gender had Californian targets the same distance from their conservative targets. Similarly, conservative (i.e., backer) GOOSE targets for each speaker were calculated as the average of the speaker's second lowest GOOSE F2 value and the average of all same-gender speakers' second lowest GOOSE values, and Californian (i.e., fronter) GOOSE targets were calculated by adding the average of same-gender speakers' ranges of GOOSE production minus 100 Hz (females: 478 Hz; males: 457 Hz) to each speaker's conservative target. (This 100 Hz adjustment to the Californian GOOSE targets was implemented because early testing of the manipulation process revealed that, without the adjustment, some preliminary manipulated Californian GOOSE tokens were so fronted as to sound unnatural.) Finally, all but four GOOSE tokens followed coronals, a phonetic environment

that favors fronting by 244 Hz on average (Labov 2010:264), so GOOSE F2 targets were adjusted for *removed* and *food* by −244 Hz and for *room* by −122 Hz. For all twelve speakers, including the two speakers whose targets are depicted in Figure 2, conservative GOOSE targets were further back than conservative TRAP targets, while Californian TRAP targets were further back than Californian GOOSE targets.

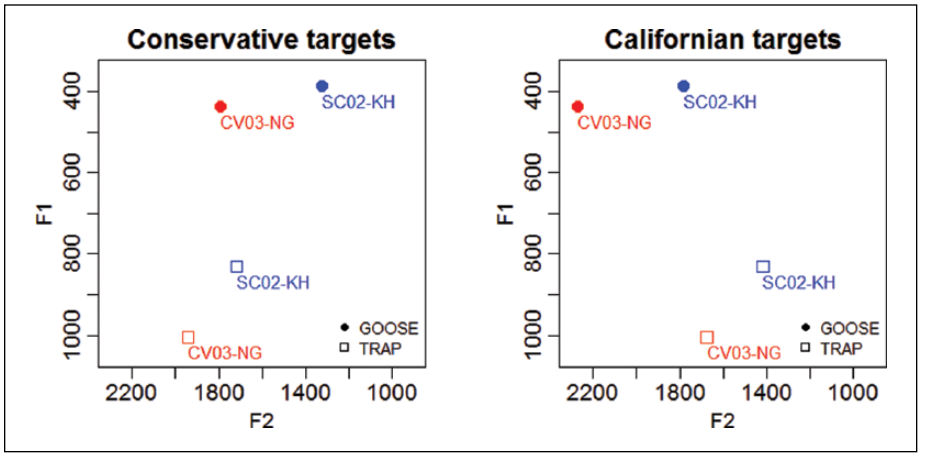

Figure 2. Conservative and Californian Targets for Speakers CV03-NG (Female) and SC02-KH (Male)
Note: Plotted F1 values represent each speaker's average F1 for that vowel; F1 was not manipulated.

This study augmented the basic process for vowel resynthesis (Styler 2017) with several adaptations necessary to preserve vowels' naturalness. These adaptations included procedures to adjust formants iteratively, match the manipulated intensity contour to that of the original, and smooth discontinuous formant transitions with neighboring sonorants. (Finer details of the manipulation process are omitted here for space, but Villarreal 2016a includes a more detailed description of the target calculation and manipulation processes.) Praat scripts (freely available at https://github.com/djvill/Vowel-Manipulation) produced matched guises from each of the twenty-four excerpts (together including ninety-one TRAP and sixty-one GOOSE tokens). Figure 3 demonstrates the end result of this process on the level of an individual TRAP token.

It is generally not possible to manipulate vowel formants to the exact Hz value of a given target; for example, speaker BA05-TN had a Californian TRAP F2 target of 1439 Hz, while the manipulated Californian token of *back* in Figure 3 had an F2 value of 1449 Hz, making the manipulated token 10 Hz off target. This difference is less than one 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), defined as 33.09 Hz for TRAP F2 and 21.86 Hz for GOOSE F2 (Kewley-Port & Watson 1994:492); in other words, listeners cannot perceive a difference between *back* tokens with F2 values of 1439 Hz versus 1449 Hz. All 182

manipulated TRAP tokens' F2 values were within one JND of the target; manipulated TRAP tokens were 11.20 Hz off target on average. All but seven of the 122 manipulated GOOSE tokens' F2 values were within one JND of the target; manipulated GOOSE tokens were 8.31 Hz off target on average, with the maximum off-target token 38.25 Hz off target (it is not clear what made this token so off-target). In other words, there was no perceptible difference between manipulated tokens and their targets. As a final check, a trained phonetician listened to all forty-eight manipulated stimuli and judged that they sounded satisfactorily natural.

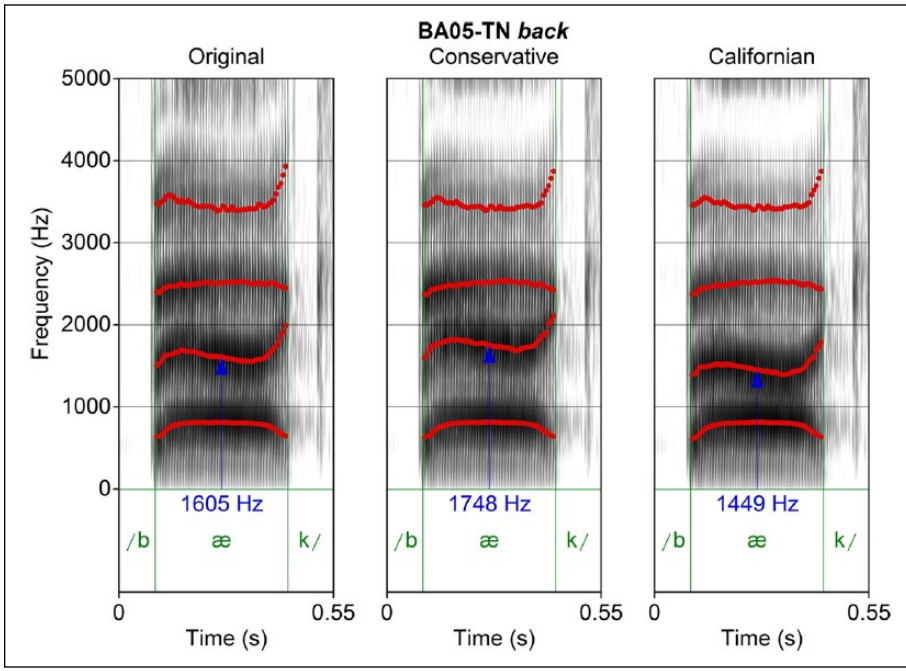

Figure 3. Spectrograms and Formant Tracks of Original, Conservative, and Californian Versions of the Token *Back* by Speaker BA05-TN
Note: Formant tracks are in light gray. Arrows indicate F2 value at the vowel midpoint.

3.2. Perceptual Task Design

The perceptual task was conducted via an online survey, the ostensible purpose of which was "evaluat[ing] candidates for a local radio show job in California that involves story telling." This "radio job" framing (modeled on that of Labov et al. 2011) provided context for otherwise decontextualized questions about a speaker's personal characteristics and encouraged listeners to candidly share their attitudes, positive or negative. (Listeners were debriefed at the end of the survey.) There were forty-eight possible

trials in the survey, one for each stimulus, arranged into twelve trial groups, one for each speaker. Each listener completed six trials, as the survey software, Qualtrics, randomly chose one trial apiece from six randomly chosen trial groups; as a result, no listeners heard more than one stimulus from any speaker.

In each trial, listeners heard a stimulus, identified the speaker's region, and rated the speaker on twelve affective scales. Scales responses were recorded via a continuous slider bar whose position was converted to an integer between 0–100. Listeners were also asked to provide narrative explanations for their regional identification and scales responses. The survey also collected listener variables such as listener region, listener age, and student status, and the effect of these variables was also analyzed.

The possible regions for the regional identification question were the Bay Area, Lower Central Valley, Southern California, and "outside California" (with a write-in blank). This question was accompanied by a version of the black-and-white California map used by Bucholtz et al. (2007) with the same regional outlines as in Figure 1. The twelve affective scales were the following: *feminine–masculine*, *Californian–not Californian*, *fast–slow*, *young–old*, *confident–not confident*, *relaxed–excited*, *friendly–not friendly*, *sounds like a Valley girl–doesn't sound like a Valley girl*, *rich–poor*, *familiar–unfamiliar*, *sounds like me–doesn't sound like me*, and *suitable–not suitable* (for a job requiring speaking to an audience). Several factors motivated the inclusion of these scales: frequency of responses in the pretesting task (see 3.1), previous research on the social meanings of the CVS (Eckert 2008b; Podesva 2011; D'Onofrio 2015), and complementing the measurement of listeners' attitudes with status and self-comparison scales.

Listeners were recruited in spring 2015 via friend-of-friend sampling through contacts across the state. All listeners were from California, self-identified as Californians, and spoke English proficiently. All but fourteen listeners were from the Bay Area (26), the Lower Central Valley (31), or Southern California (26). This data set consists of 580 trials from ninety-seven listeners, as ninety-five listeners responded to six trials and two responded to five trials (due to audio failure).

3.3. Bayesian Data Analysis

Survey data were analyzed via Bayesian models, which calculate posterior distributions (see below) that assign different levels of credibility to possible values of given parameters (Kruschke 2015). To my knowledge, Bayesian methods have not been previously applied to data in perceptual sociolinguistics, despite the fact that Bayesian inference has several advantages over the methods more commonly used (frequentist inference) to conduct inferential statistics; for example, Bayesian models avoid issues with convergence (Kimball et al. forthcoming) and produce more accurate estimates of effects (Eager & Roy 2017). Bayesian inference is not without its disadvantages: it has traditionally been relatively inaccessible, Bayesian analyses can be complex and time-consuming to set up, and these analyses require extensive computing power and coding acumen.

Bayesian inference also differs from frequentist inference in its end product; whereas frequentist methods produce a p value, a probability of observed data given a

null hypothesis about the value of a given parameter, Bayesian methods produce a "posterior distribution," which assigns different levels of credibility to possible values of given parameters. The posterior distribution is calculated from the product of the "prior distribution" (which describes 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 para-meters) divided by the sum or integral of this product across all values of the parameters. This formula is impossible to solve directly for all but the most elementary applications,[3] so for more complex models, the posterior distribution is instead approximated via Markov chain Monte Carlo (MCMC) probabilistic sampling. This method generates a sample of posterior values by walking a "chain" through possible values for the posterior, the distribution of which has its mode at the most credible posterior value for that parameter. With a sufficiently long MCMC chain—conventionally defined as an effective sample size (ESS) of 10,000 chain steps—the distribution of values visited by the chain approximates the "true" posterior distribution with tolerably low error (Kruschke 2015).

Another common Bayesian 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 effects' magnitudes (see, e.g., Gelman et al. 2013:95). While I do consider effects' magnitudes in the following discussion, I also find it useful to differentiate effects that are credibly meaningful from those that may be due to chance; at the same time, this approach raises the possibility of Type I error due to making multiple comparisons. Similar to the frequentist practice of testing significance against a null hypothesis of zero using a frequentist *p* value or confidence interval, this study used a Bayesian decision rule that worked as follows. MCMC sampling generates estimates for different parameters at each step in the MCMC chain; it is thus possible to calculate, at each step, the difference between estimates for two parameter values (say, level A versus level B of factor 1), and then to treat the distribution of these differences as a marginal posterior distribution, a "posterior contrast." From this posterior contrast, a 95 percent highest density interval (HDI) can be calculated, an interval that includes the most credible 95 percent of possible differences between level A and level B. (Readers familiar with frequentist confidence intervals can consider these to be roughly analogous to HDIs.) This situation leads to two possible outcomes: if the 95 percent HDI of a posterior contrast between level A and level B of factor 1 does not include zero, then zero is not a credible value for the difference between responses to level A versus level B; the difference between level A and level B is thus deemed to be a "credible difference" and the effect of factor 1 is deemed to be a "credible effect." Conversely, if the 95 percent HDI of the posterior contrast does include zero, then zero is a credible value for the difference between responses; the effect is thus not deemed to be credible—though this is not tantamount to an affirmative statement about the lack of an effect (Kruschke 2015). This difference in approach to statistical inference is summarized in (1).

$$(1)\quad \text{In}\left\{\begin{array}{l}\text{frequentist inference}\\ \text{Bayesian inference}\end{array}\right\},\ \text{a difference/effect is said to be}\left\{\begin{array}{l}\text{significant}\\ \text{credible}\end{array}\right\}\ \text{if ...}$$
$$\left\{\begin{array}{l}\text{the } p \text{ value of observed data given a null hypothesis of 0 is less than } \alpha\ (\text{e.g., .05})\\ \text{an HDI (e.g., at 95\%) of the posterior contrast between 2 parameters excludes 0}\end{array}\right\}.$$

The models used in this analysis were "hierarchical models," meaning that the parameters used to compute likelihood functions are given prior distributions whose parameters ("hyperparameters") are not fixed but themselves are estimated via higher-level prior distributions. These models were run using R (R Core Team 2015) and the Bayesian sampling program JAGS (Plummer 2003), based on R scripts from Kruschke (2015); these scripts are available from the present author upon request. For each simulation, diagnostic plots for selected parameters and hyperparameters were 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. The details of these models are given in their respective subsections below.

4. Results

This section describes the results of this task in terms of the two types of data that the survey yielded: regional identification data (map question) and affective scales data. In the analysis of both regional identification and scales data, the main predictor of interest was GUISE: did listeners' perceptions change based on whether they heard California-shifted vowels or non-shifted vowels? The analysis of scales data reported here also includes speaker-dependent factors (SPEAKER REGION, SPEAKER GENDER, and SPEAKER ETHNICITY), interactions of these factors, and listener effects. I also briefly describe the Bayesian models by which the analysis was carried out.

4.1. Regional Identification

Across all 580 trials, speakers were identified as from the Bay Area in 24.5 percent of the trials, the Lower Central Valley in 25.3 percent of the trials, Southern California in 32.8 percent of the trials, and outside California in 17.4 percent of the trials. (These responses serve as a predictor in the analysis of scales discussed in 4.2.) Among the 479 trials in which listeners identified speakers as being from the Bay Area, Lower Central Valley, or Southern California, listeners correctly identified speakers' region in 172 trials, or 35.9 percent. The posterior distribution of estimates of accurate-recognition rates,[4] with a mode of 35.9 percent and 95 percent HDI limits at 31.7 percent and 40.2 percent, does not exclude 33.3 percent (the level of chance guessing) as a credible value, indicating a lack of credible evidence that listeners' true rate of accurate recognition is greater than chance. 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.

The analysis of regional identification responses was carried out via Bayesian hierarchical models in which distributions of responses in each category were modeled

with a multinomial distribution as the likelihood function and Dirichlet distribution priors (Gelman et al. 2013). The priors for the individual proportions in the Dirichlet distributions were beta distributions with shape parameters 2 and 6 (representing a noncommittal prior assumption that the identification rate for each cell was 25 percent). The same one-predictor hierarchical model structure was used for all predictors individually, although only the analysis of GUISE is reported here.

Figure 4 displays the results of this analysis: the conservative minus Californian posterior contrasts for each response category, in units of proportions (i.e., proportion conservative responses minus proportion Californian responses). Whereas GUISE had no credible effect on the rate at which listeners identified speakers as being from the Bay Area or Lower Central Valley, GUISE did credibly affect identification rates for Southern California and outside California; as Figure 4 shows, for Southern California the entire 95 percent HDI of possible conservative-minus-Californian differences is less than zero, and for outside California the entire 95 percent HDI is greater than zero, excluding zero as a credible conservative-minus-Californian difference for either response category. In other words, regional identification was sensitive to the formant manipulation, as listeners associated California-shifted TRAP/GOOSE with Southern California and non-shifted TRAP/GOOSE with non-Californians. (The models used to analyze regional identification did not permit an assessment of interactions, but descriptive statistics suggest that these main effects of GUISE also apply to eleven of twelve individual speakers.) This result suggests that Californians do perceive these CVS features as Californian, and specifically as a Southern Californian phenomenon.

4.2. Scales

Scales data were standardized to control for listeners' differential use of the slider bars, as the average listener used a range of 92.8 and some listeners used far less. All results are reported here in unit-less standard deviations (i.e., *z*-scores). Across all listeners, the mean rating was 55.3 and the standard deviation was 25.1; a difference of one standard deviation can thus be interpreted as a difference of roughly one quarter of the slider bar for the average listener in this survey.

The analysis of scales data reported here was carried out via eight-predictor hierarchical models—with all twelve scales modeled separately using identical models—that included the primary predictors (GUISE, SPEAKER REGION, SPEAKER GENDER, and SPEAKER ETHNICITY), LISTENER REGION, and PERCEIVED REGION (i.e., listeners' regional identification of speakers); LISTENER and SPEAKER were included as random effects. Second-order predictors were also included to assess interaction effects.[5] Scales ratings were modeled with a normal distribution as the likelihood function and various priors on the hyperparameters of this likelihood function. At each step in the chain, this model calculated 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) as a means of assessing the effects of different predictors.

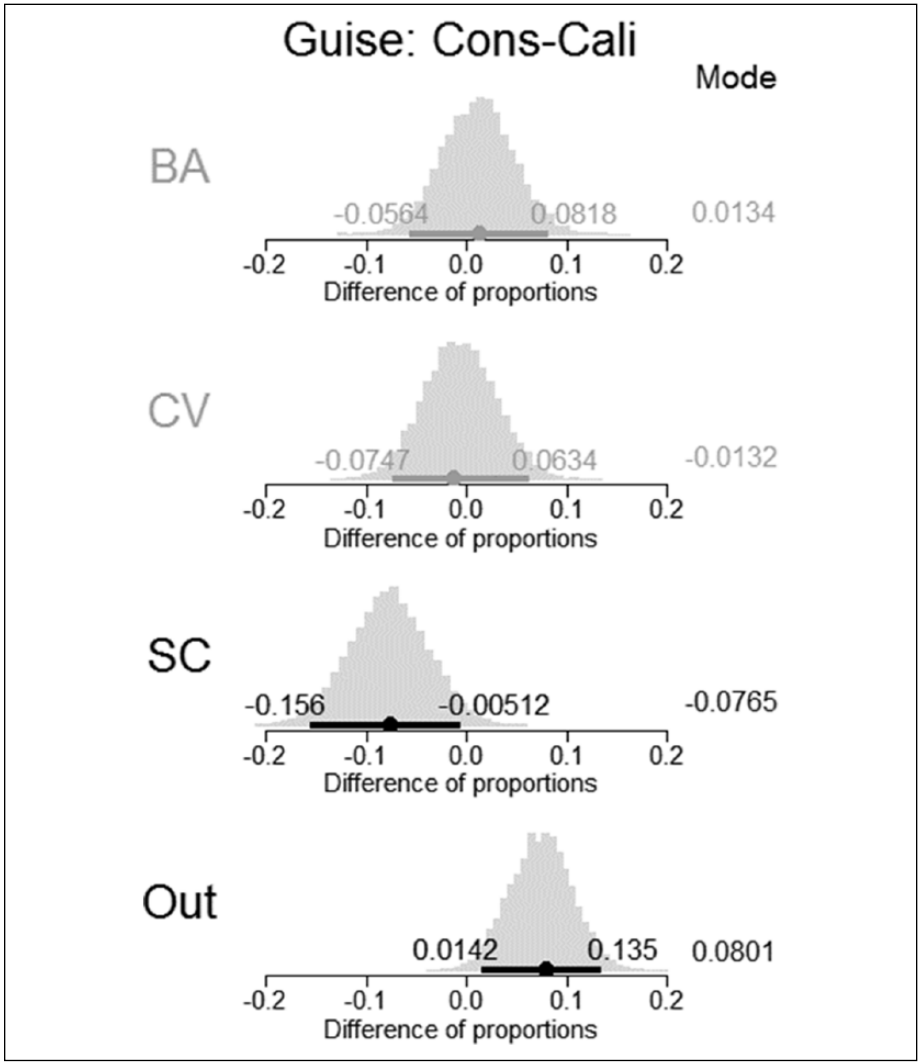

Figure 4. Histograms of Posterior Regional Identification Contrasts by Guise: Cons[ervative] Minus Cali[fornian]

Note: Regions are Bay Area (BA), Lower Central Valley (CV), Southern California (SC), and outside California (Out). Bars indicate 95 percent HDIs; dots indicate modes. Black bars/labels indicate response categories with contrasts credibly above or below zero; gray bars/labels indicate contrasts not credibly above or below zero.

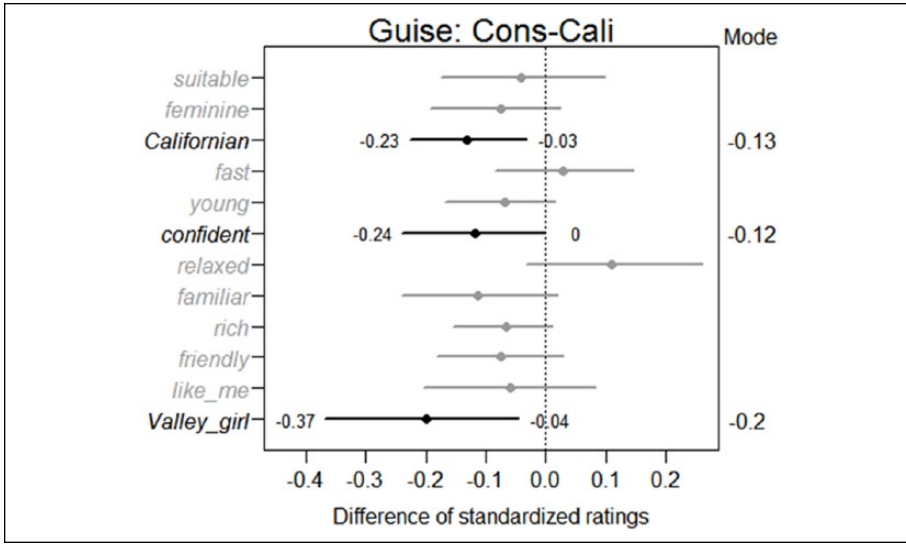

Figure 5. Modes and 95 Percent HDIs of Posterior Scales Contrasts by Guise: Cons[servative] Minus Cali[fornian]
Note: Bars indicate 95 percent HDIs; dots indicate modes. Black bars indicate scales with contrasts credibly above or below zero; gray bars indicate contrasts not credibly above or below zero. The upper bound for the *confident* 95 percent HDI is actually −0.0036, but appears as zero due to rounding error.

4.2.1. Main Effects. GUISE credibly affected responses to three scales, *Californian*, *confident*, and *sounds like a Valley girl*, with higher ratings for California-shifted stimuli on each. The posterior contrasts are displayed in Figure 5, which can be thought of as a birds-eye view of the posterior contrast histograms in Figure 4 (excluding the tails of the distributions beyond the 95 percent HDI limits). All of these credible differences are small in magnitude; the largest credible difference for *Californian*, for example, is a 0.23 standard deviation increase for the Californian guise above the conservative guise, corresponding to a difference of about six points on the 0–100 slider bar.

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 6 indicates, the higher ratings for the Californian guise are relative to different baselines: a high baseline rating for *Californian* and a low baseline rating 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.

SPEAKER REGION (i.e., the region that speakers were actually from) had a credible effect on responses to eight scales: all but *Californian*, *relaxed*, *familiar*, and *sounds like me*. 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*.

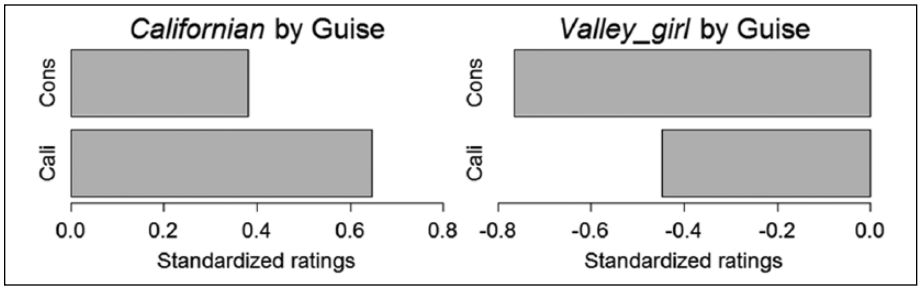

Figure 6. Mean Standardized *Californian* and *sounds like a Valley girl* Ratings by Guise (Cali[fornian] and Cons[ervative])

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* (mode: 2.07) and *Valley girl* (mode: 0.71) than any other primary predictor had on any other scales. This result is not entirely surprising, as these two scales are inherently gendered, and speakers' gender was transparent to listeners in a way that guise, speaker region, or ethnicity were not.

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*. Paradoxically, the only scale for which ethnicity did not credibly affect ratings was *sounds like a Valley girl*, despite the fact that the Valley girl character (and other folk depictions of California English speakers) originally represented a Caucasian stereotype, meaning this scale comes closer to being inherently raced than any other.

Table 1 summarizes these results by listing the credible effects of each primary predictor: guise, speaker region, speaker gender, and speaker ethnicity. As measured by the number of scales credibly affected by each predictor, with credible differences on just three scales, guise had a weaker effect than the other primary predictors. (Of course, these scales are by no means representative of all of the responses that Californian listeners may have to language variation, so a count of credible effects is merely a rough proxy for the strength of each predictor's effect on listeners' perceptions.) This small effect of guise is also clear from the posterior contrasts of the three credible differences by guise, as the models do not rule out the effect of guise being as small as 1.25 notches on the 0–100 slider bar (0.05 standard deviations). However, whereas other predictors had a larger effect on *confident* and *Valley girl* than guise did,

the effect of GUISE on *Californian* was stronger than other predictors' effect on *Californian*; in other words, these listeners associated Californianness more strongly with California-shifted TRAP/GOOSE than with speakers' region, gender, or ethnicity. 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, except when it comes to speakers' Californianness. Speakers' ethnicity, by contrast, registers rather strongly in listeners' ratings.

4.2.2. Interactions of Primary Predictors. Statements about the credibility of interaction effects are based on the same decision rule for main effects: interaction effects are deemed to be credible if the 95 percent HDI of the posterior interaction contrast excludes zero.

Table 1. Credible Effects by Primary Predictor and Scale

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

Note: X > Y indicates credibly higher rating for X than Y. Gray cells indicate no credible effect. Gender: F[emale] vs. M[ale]. Ethnicity: Latina/o vs. Cauc[asian].

The GUISE × SPEAKER REGION and GUISE × SPEAKER ETHNICITY interactions yielded no credible effects. In other words, the difference between conservative and Californian guises applied equally to speakers from all three regions; the difference in scales ratings between guises applied equally to both Latina/o and Caucasian speakers, regardless of whether GUISE or SPEAKER ETHNICITY had a credible main effect on ratings. The GUISE × SPEAKER GENDER interaction yielded only one credible difference: *confident*; 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 × SPEAKER GENDER interaction effects. That is, not only was the Californian guise rated more *Californian* and *Valley girl* than the conservative guise, this GUISE effect was virtually the same for both female and male speakers. Likewise, the credible SPEAKER GENDER effects on *suitable*, *feminine*, *Californian*, *fast*, *relaxed*, *friendly*, and *Valley girl* were virtually the same for both conservative and Californian guises. Whereas California-shifted stimuli were rated higher on *Californian* and *Valley girl* than non-shifted stimuli and females were rated higher than males, these effects did not compound one another.

The interaction effects of GUISE with other primary predictors thus represent additional evidence that speakers' use of CVS-like TRAP/GOOSE registers weakly in Californian listeners' overall responses to language variation; save for the interaction effect with SPEAKER 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 suggests 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 interaction between SPEAKER GENDER and SPEAKER ETHNICITY yielded five credible differences: *feminine, fast, confident, relaxed*, and *friendly*. Each of these scales also had independent credible main effects for both SPEAKER GENDER and SPEAKER ETHNICITY, with female speakers patterning with Latina/o speakers and male speakers patterning with Caucasian speakers in each case; male Caucasian speakers were rated substantially different from female Latinas, female Caucasians, or male Latinos. The SPEAKER REGION × SPEAKER GENDER and SPEAKER REGION × SPEAKER ETHNICITY interactions also indicate a high degree of interrelation between these three predictors; the SPEAKER REGION × SPEAKER GENDER interactions yielded credible effects on all twelve scales.

To summarize the results of the interaction analysis, 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, shifted TRAP/GOOSE 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 listener perceptions—but it is real, as the main effects of GUISE are not simply a function of acoustic differences in TRAP/GOOSE F2 piggybacking on other speaker characteristics (although as I discuss in 4.2.3, in some cases speaker characteristics may have led to a reversal of the overall GUISE effect). The interactions not involving GUISE, by contrast, indicate a high degree of interrelation between speaker region, gender, and ethnicity. The SPEAKER GENDER × SPEAKER ETHNICITY interactions showed that male Caucasian speakers stand apart from other speakers on *feminine*, *fast*, *confident*, *relaxed*, and *friendly*, suggesting that maleness and Whiteness compound the attitudinal effects of one another in the ears of Californian listeners.

4.2.3. Listener and Speaker Effects. Overall, there were few effects of listeners' characteristics on scales ratings. In terms of LISTENER REGION, listeners from the Lower Central Valley rated speakers as more *confident* overall than did Southern California listeners, and less *young* overall than did both Bay Area and Southern California listeners. LISTENER REGION also yielded one credible interaction with GUISE, as the Californian guise was rated less *feminine* by Bay Area listeners and more *feminine* by Southern California listeners. That is, whereas Bay Area listeners associate California-shifted TRAP/GOOSE with masculinity, Southern California listeners associate these CVS features with femininity. The reasons for this regional difference in gendered social meanings of CVS-like TRAP/GOOSE remain unclear to me, especially given the feminization of the CVS in popular portrayals (Eckert 2008b). Aside from the scale *feminine*, however, the CVS received the same overall evaluations by listeners from across the state.

Beyond LISTENER REGION (which was included in the eight-predictor hierarchical models discussed in 4.2), the effects of numerous listener effects were analyzed via individual two-predictor hierarchical models that included SPEAKER as a random effect; as with the eight-predictor models discussed above, scales ratings were modeled with a normal distribution as the likelihood function and various priors on the hyperparameters of this likelihood function. Just two factors emerged from these models as having at least one credible effect on scales ratings: LISTENER MOBILITY (the extent to which listeners had lived in different places) and LISTENER AGE. LISTENER MOBILITY affected ratings of *suitable* and *young*, but neither set of credible differences suggested a clear cumulative effect of increased mobility on listeners' ratings; rather, it is likely that the credible effects of LISTENER MOBILITY on *young* simply reflected the preponderance of Lower Central Valley listeners in the less-mobile categories. In addition, listeners aged forty or older rated speakers across the board as more *young* than did listeners younger than thirty. The other listener-dependent factors collected in the survey yielded no credible effect on scales responses: LISTENER GENDER, LISTENER ETHNICITY, LISTENER MONO/BI/MULTILINGUALISM, LISTENER EXPERIENCE WITH LINGUISTICS, and STUDENT STATUS.

There were several credible effects of individual speakers on scales results beyond those already captured by main effects of SPEAKER REGION, SPEAKER GENDER, and/or SPEAKER ETHNICITY. For example, SPEAKER credibly affected *sounds like a Valley girl*, as, for example, female speaker SC06-KY was rated lower on *sounds like a Valley girl* than male speaker CV04-BW. Although the interaction of SPEAKER and GUISE was excluded from the inferential model, descriptive statistics by SPEAKER and GUISE reveal that two speakers' conservative guises were rated higher on *sounds like a Valley girl* than their Californian guises, counter to the general relationship in Figure 6, suggesting that something about the speech of these two speakers resulted in atypical perceptions of their CVS use. These results do not indicate that the main effect of GUISE on *sounds like a Valley girl* is really the result of idiosyncratic speaker characteristics, however, as the majority of speakers were rated higher on *sounds like a Valley girl* in the Californian guise. Seven scales exhibited no credible speaker effects, including *Californian* and *confident*, the other two scales (along with *sounds like a Valley girl*) with a credible main effect of GUISE.

5. Discussion

To return to the first research question, both the regional identification and affective scales data in this study suggest that Californian listeners do recognize a contrast between CVS and non-CVS vowels, represented here by TRAP/GOOSE. This finding compares to that of Soukup's (2013) study of listener uptake of the social meanings of *Dialekt*, with an important difference: the credible effects of GUISE in the present study reveal that the contrast between CVS and non-CVS vowels is perceptually salient to Californian listeners on a less overt level, as compared to Soukup's (2013) listeners' overt recognition of a contrast between standard and dialectal German. This result indicates that listeners' recognition of the contrastiveness of forms, like other forms of folk-linguistic awareness, can be categorized along a dimension of availability from overt to implicit (Preston 1996).

To return to the second research question, CVS features are associated among Californian listeners not only with sounding more Californian and more like a Valley girl, but also more confident. The effect of these features on the perception of confidence is mitigated by speakers' gender; men who use California-shifted TRAP/GOOSE sound more confident than those who do not, whereas women sound equally confident regardless of TRAP/GOOSE production. This result stands in contrast to discourses that attribute a lack of confidence to Californian 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). While folk-linguistic discussions of Californian speech that foreground high rising terminals stereotype Californian speech as non-confident, this association with a lack of confidence does not spread to these CVS features; on the contrary, male speakers' use of California-shifted TRAP/GOOSE is interpreted as a sign of confidence.

While SPEAKER GENDER, SPEAKER REGION, and SPEAKER ETHNICITY had sizeable effects on scales ratings, speakers' use of CVS-like TRAP/GOOSE had a real but small effect; as compared to GUISE, the speaker variables of GENDER, REGION, and ETHNICITY yielded not only a greater number of credible effects, but also effects of greater magnitude. GUISE also yielded almost no credible interaction effects with SPEAKER GENDER, SPEAKER REGION, or SPEAKER ETHNICITY. These results suggest that when Californians make judgments about speakers' personal traits, they rely primarily on characteristics such as speakers' gender and ethnicity, with TRAP/GOOSE playing only a secondary role and affecting listeners' judgments to a more limited extent. Whereas the effects of speakers' gender, region, and ethnicity interact with one another in influencing listeners' judgments, however, listeners appear to judge speakers' use of these CVS features mostly independently of other speaker characteristics; this result lends credence to the main effects of guise found in this study, as it suggests that interspeaker differences in speech were not responsible for the main effects of guise.

6. Revisiting Social Meaning

This research indicates that while Californian listeners are sensitive to features of the CVS, recognizing the contrast between California-shifted and non-shifted TRAP and

GOOSE, their perceptions of two CVS features differ from the social meanings suggested by previous production research: carefreeness (Podesva 2011), an opposition to gang affiliation (Fought 1999), and Whiteness, femininity, and privilege (Eckert 2008b). While the results of the present study show that listeners associate California-shifted TRAP and GOOSE with confidence (for males), Californianness, and sounding like a Valley girl, these results do not support a direct association between these CVS features and the perception of relaxedness, wealth, or femininity—at least not in this particular context.

Theories of social meaning indicate at least two ways to account for this mismatch between social meanings. First, in line with the idea that we cannot uncritically base claims about social meaning on production alone (Schilling 2013), it is possible that while these social meanings are suggested by speakers' stylistic use of the CVS, these same meanings simply do not exist in perception. This account stands as a useful caution against overstating the role of speaker agency in transmitting social meanings, but its lack of predictive power renders it theoretically unsatisfying; if accepted, this account could justify uncritically nullifying any production-based claims about social meaning. A second, more interesting account of this social-meaning mismatch rests on the notion that the interpretation of linguistic forms is shaped in interaction by the contextualization cues present in the interaction (Gumperz 1982; Soukup 2013)—an account that helps to explain how such diverse meanings as "nerd girl," "Orthodox Jewish boy," and "gay diva" (Eckert 2008a:469) can occupy the same indexical field. That is, if listeners are to make sense of the social meanings present in production, then contextualization cues are necessary, as these cues "slice" the indexical field into a subset of possible meanings relevant to the current interaction, effectively guiding the process of listener interpretation (Campbell-Kibler 2007; Pharao et al. 2014). To summarize this interpretation, the absence of relaxedness, wealth, or femininity in listener perceptions of California-shifted TRAP/GOOSE does not necessarily falsify the potential existence of these meanings in the CVS's indexical field, but leaves open a second possibility: these meanings do exist in the CVS's indexical field but did not emerge in this particular task due to the absence of the contextualization cues that would have activated these meanings in listener perceptions.[6]

The current study leaves open several avenues for improvement and future study. First, the fact that TRAP and GOOSE together elicit perceptions tied to California begs investigation of how the rest of the CVS factors into perceptions of Californianness; if non-manipulated CVS variables like KIT and GOAT that co-occurred with manipulated TRAP/GOOSE influenced listeners' interpretation of the non-shifted vs. California-shifted vowels, this would lend credence to the contextualization-cues account discussed above. Similarly, the tandem manipulation of TRAP and GOOSE in this study indicates the need for a detangling of these features' perceptual effects. Second, this study leaves underexplored the roles of gender and ethnicity in the social meanings of the CVS, even though these findings suggest that speakers' gender and ethnicity shape listeners' perceptions of speakers to a greater degree than the position of two CVS features. Third, given that the CVS is hardly unique to California, it is an open question whether the CVS indexes California outside the state or takes on locally salient values. These

open questions notwithstanding, the current study's attempt to square social meanings as they manifest in production and perception represents progress in our understanding of the construction of social meaning.

Appendix A

Text of Excerpts

Following is the text of the twenty-four excerpts (two excerpts apiece from twelve 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; all 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 I dunno 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

(continued)

Appendix. (continued)

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

Acknowledgments

I would like to acknowledge the support of my doctoral advisor, Robert Bayley, and committee members: Vai Ramanathan, Valerie Fridland, and Santiago Barreda. Thanks are also due to Christina Calvillo, Michael Shepherd, Dave Corina, Natalie Operstein, Andrew Wong, Julie Ngo, Chris Eager, and anonymous reviewers.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research received financial support from two units at the University of California, Davis: the Graduate Group in Linguistics and the Office of Graduate Studies.

Notes

1. This paper uses Wells' (1982) notation for vocalic lexical sets, as this system does not presuppose static articulatory locations implied by IPA symbols. TRAP stands in for the phoneme canonically represented as /æ/ (with TRAP-N for prenasal TRAP), GOOSE for /u/, KIT for /ɪ/, and GOAT for /o/.
2. Here, I use Agha's (2003:231) definition of "enregisterment" as "processes through which a linguistic repertoire (or bundle of features) becomes differentiable within a language as a socially recognized register of forms." Thanks to an anonymous reviewer for making the point about the status of TRAP and GOOSE vis-a-vis the enregisterment of California English.
3. Thanks to Chris Eager for clarifying this point for me.
4. 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 true accurate-recognition rate is 33.3 percent) and a Bernoulli likelihood function with 172 successes in a sample size of 479.
5. Unlike ANOVA and other frequentist models, Bayesian models need not include any interaction terms beyond that which the experimenter seeks to analyze (Kruschke 2015).
6. Moreover, these two accounts of mismatches between social meanings in production and in perception are not totally mutually exclusive. It is possible that some of the social meanings of the CVS suggested by production studies could emerge in perception given the right contextualization cues—but that some of these meanings simply cannot be found in perception at all. Indeed, we must avoid viewing contextualization cues as unlocking boundless possibilities for speakers to transmit social meaning, as this idea wrongly affords total agency over meaning to speakers and imposes no structure on possible meanings (Schilling 2013).

ORCID iD

Dan Villarreal http://orcid.org/0000-0002-6070-1138

References

Agha, Asif. 2003. The social life of cultural value. *Language & Communication* 23(3-4). 231-273.

Boersma, Paul & David Weenink. 2015. Praat [phonetic analysis software]. Version 5.4.01, retrieved Nov 2014 from http://www.fon.hum.uva.nl/praat/

Bucholtz, Mary, Nancy Bermudez, Victor Fung, Lisa Edwards & Rosalva Vargas. 2007. Hella Nor Cal or totally So Cal? The perceptual dialectology of California. *Journal of English Linguistics* 35(4). 325-352.

Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech* 82(1). 32-64.

Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21(1). 135-156.

D'Onofrio, Annette. 2015. Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics* 19(2). 241-256.

D'Onofrio, Annette, Penelope Eckert, Robert J. Podesva, Teresa Pratt & Janneke Van Hofwegen. 2016. The low vowels in California's Central Valley. In Valerie Fridland, Betsy Evans, Tyler Kendall & Alicia Wassink (eds.), *Speech in the western states*, vol. 1, 11-32. 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(1). 60-80.

Eager, Christopher D. & Joseph Roy. 2017. Mixed models are sometimes terrible. Paper presented at the Linguistic Society of America, Austin, TX.

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

Eckert, Penelope. 2008a. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453-476.

Eckert, Penelope. 2008b. Where do ethnolects stop? *International Journal of Bilingualism* 12(1-2). 25-42.

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

Fought, Carmen. 1999. A majority sound change in a minority community: /u/-fronting in Chicano English. *Journal of Sociolinguistics* 3(1). 5-23.

Fridland, Valerie, Kathryn Bartlett & 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). 1-16.

Fridland, Valerie, Tyler Kendall, Betsy Evans & Alicia Wassink (eds.). 2016. *Speech in the western states*, vol. 1. Durham, NC: Duke University Press.

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

Gumperz, John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Kennedy, Robert & James Grama. 2012. Chain shifting and centralization in California vowels: An acoustic analysis. *American Speech* 87(1). 39-56.

Kewley-Port, Diane & Charles S. Watson. 1994. Formant-frequency discrimination for isolated English vowels. *Journal of the Acoustical Society of America* 95(485). 485-496.

Kimball, Amelia E., Kailen Shantz, Christopher D. Eager & Joseph Roy. Forthcoming. Beyond maximal random effects for logistic regression: Moving past convergence problems.

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

Labov, William. 2010. *Principles of linguistic change*, vol. 3, *Cognitive and cultural factors*. Hoboken, NJ: Wiley-Blackwell.

Labov, William, Sharon Ash & Charles Boberg. 2006. *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 & Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15(4). 431-463.

Montello, Daniel R., Alinda Friedman & Daniel W. Phillips. 2014. Vague cognitive regions in geography and geographic information science. *International Journal of Geographical Information Science* 28(9). 1802-1820.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1). 62-85.

Pharao, Nicolai, Marie Maegaard, Janus Spindler Møller & Tore Kristiansen. 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43(1). 1-31.

Plummer, Martyn. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling [Computer program]. Version 4.0.0, retrieved Oct 2015 from http://mcmc-jags.sourceforge.net/

Podesva, Robert J. 2011. The California Vowel Shift and gay identity. *American Speech* 86(1). 32-51.

Pratt, Teresa & Annette D'Onofrio. 2014. Awareness in enregisterment: Performance of the California Vowel Shift in SNL's "The Californians." Paper presented at NWAV 43, Chicago, IL.

Preston, Dennis R. 1996. Whaddayaknow? The modes of folk linguistic awareness. *Language Awareness* 5(1). 40-74.

Preston, Dennis R. 2011. Language, people, salience, space: Perceptual dialectology and language regard. *Dialectologia* 5. 87-131.

R Core Team. 2015. R: A language and environment for statistical computing [Computer program]. Version 3.2.2, retrieved Aug 2015 from https://www.R-project.org/

Schilling, Natalie. 2013. Investigating stylistic variation. In J. K. Chambers & Natalie Schilling (eds.), *The handbook of language variation and change*, 327-349. Oxford: Wiley Blackwell.

Soukup, Barbara. 2013. Austrian dialect as a metonymic device: A cognitive sociolinguistic investigation of Speaker Design and its perceptual implications. *Journal of Pragmatics* 52. 72-82.

Styler, Will. 2017. Using Praat for linguistic research [Online resource]. Version 1.7.2, retrieved Nov 2017 from http://savethevowels.org/praat/UsingPraatforLinguisticResearchLatest.pdf

Tyler, Joseph C. 2015. Expanding and mapping the indexical field: Rising pitch, the uptalk stereotype, and perceptual variation. *Journal of English Linguistics* 43(4). 284-310.

Villarreal, Dan. 2016a. The construction of social meaning: A matched-guise investigation of the California Vowel Shift. Davis, CA: University of California, Davis, PhD dissertation.

Villarreal, Dan. 2016b. "Do I sound like a Valley Girl to you?" Perceptual dialectology and language attitudes in California. In Valerie Fridland, Tyler Kendall, Betsy Evans & Alicia Wassink (eds.), *Speech in the western states*, vol. 1, 55-75. Durham, NC: Duke University Press.

Warren, Ron & 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.

Watson, Kevin & Lynn Clark. 2013. How salient is the nurse~square merger? *English Language and Linguistics* 17(2). 297-323.

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

Williams, Angie, Peter Garrett & Nikolas Coupland. 1999. Dialect recognition. In Dennis R. Preston (ed.), *Handbook of perceptual dialectology*, vol. 1, 345-358. Amsterdam: John Benjamins.

Author Biography

Dan Villarreal is a sociolinguist whose research interests include social meaning, variation, and folk linguistics. As of 2017, he works as a postdoctoral scholar at the New Zealand Institute of Language, Brain, and Behaviour (at the University of Canterbury) on projects investigating phonological priming and dialectal variation in New Zealand English.