Controlled and automatic perceptions of a sociolinguistic marker

ANNETTE D'ONOFRIO
Northwestern University

ABSTRACT

This paper explores the relation between controlled and automatic perceptions of a sociolinguistic variable that yields no metalinguistic commentary—a *marker* (Labov, 1972). Two experiments examine links between the backed TRAP vowel and its social meanings. The first, a matched guise task, measures social evaluations of the feature in a relatively controlled, introspective task. In the second, two measures are used that access different points in online processing and different degrees of listener control: (a) lexical categorization of an ambiguous stimulus, measured by a mouse click, and (b) automatic, early responses to this ambiguous stimulus, measured by eye movements. While listeners perceptually link TRAP-backing with social information in all three measures, specific social effects differ across the measures. Findings illustrate that the task and time course of a response influence how listeners link a linguistic marker with social information, even when this sociolinguistic knowledge is below the level of conscious awareness.

The issue of awareness has played a central role in theories of linguistic variation since the earliest sociolinguistic studies (Labov, 1972; Preston, 1996). Labov's (1972) typology of sociolinguistic variables outlines three categories reflecting speaker awareness of a variable and its social correlates. Speakers show no awareness of *indicators*, implicit awareness of *markers*, such that they shift usage of such variables with attention paid to speech, and explicit awareness of *stereotypes*; their social meaning is the "overt topic of social commentary" (Labov, 1972:178). The *marker* disentangles metalinguistic commentary from listener knowledge or awareness: behavior and evaluation are affected by individuals' knowledge of a variable's social meaning, but this knowledge is not the object of conscious discussion. Work in sociolinguistic perception has demonstrated repeatedly that listeners can show sensitivity to a feature-meaning link in tasks that need not access metalinguistic awareness of that link (Babel,

2016; Campbell-Kibler, 2012; Hay, Warren, & Drager, 2005; Koops, Gentry, & Pantos, 2008; Staum Casasanto, 2008).

The tripartite indicator-marker-stereotype distinction and its relation to the progression of a sound change (Johnstone, Andrus, & Danielson, 2006; Labov, 1972) sets forth a theoretical continuum of awareness, in which a variable that arises as a sound change below the level of consciousness garners increasing amounts of awareness over time. A unidimensional continuum of awareness suggests that sociolinguistic signs that yield conscious awareness will also influence more implicit perceptions. If a *stereotype* originated as a *marker*, for example, we might expect listeners who show metalinguistic acknowledgment of the feature to also show awareness of the feature-meaning link in more implicit behavior. However, work on psychological processes of judgment and reasoning (Evans, 2006; Kahneman & Frederick, 2002; Kahneman & Tverskey, 1972) and social bias (Bargh & Chartrand, 1999; Devine, 1989; Payne, Burkley, & Stokes, 2008) has shown that associations revealed in more implicit versus more explicit processing do not neatly mirror one another. Instead, there is evidence that individuals draw on separate systems for implicit, or unconscious, versus explicit, or conscious, processing. For example, Devine (1989) showed that groups of differently prejudiced participants all exhibited stereotype-induced behavior when a racial stereotype was activated *implicitly*, via priming, but participants with lower prejudice scores showed inhibition of this response when more controlled *explicit* self-reporting was assessed.

Campbell-Kibler (2012) outlined the utility of these "dual-processing" models to the study of sociolinguistic perception. Allowing for explicit and implicit sociolinguistic knowledge systems to operate independently from one another—rather than posing them as different points along a single dimension of awareness—can help explain how speakers actively use linguistic features to do social and interactional work, and how listeners can derive social meaning from these features, without metalinguistic or introspective awareness from either party. Such a model raises questions about exactly what is meant by "implicit" versus "explicit" sociolinguistic awareness, how these notions map onto paradigms commonly used in sociolinguistic perception, and how researchers can tap into each of these types of processes to test how implicit and explicit responses may reflect or contradict one another.

Recent studies have begun to examine this relation directly (Babel, 2016; Campbell-Kibler, 2012; Levon & Fox, 2014), indicating that explicit and implicit responses to the same sociolinguistic variables do not necessarily mirror one another. Campbell-Kibler (2012) showed that while the same associations between linguistic features and social information were visible in three tasks that elicited responses at different levels of implicitness, little correlation emerged among individuals' performance in each. Levon and Fox (2014) found that the degree of community-level metalinguistic discussion of phonological variables in British English had almost no impact on listener evaluations in a more implicit matched-guise task. In a self-paced reading task, Squires (2016) showed that implicit, online listener behavior in response to nonstandard and uncommon

morphosyntactic structures differed depending on whether a listener metalinguistically commented on these structures, though even those who did not explicitly comment showed evidence of implicitly perceiving the manipulation. These studies offer support for distinct representations and/or processing mechanisms for more explicit and more implicit sociolinguistic information.

In these studies, *explicit* knowledge of a sociolinguistic feature is extrapolated from an individual's ability to consciously report on the feature, with *implicit* knowledge being defined as an effect on some aspect of perceptual behavior below the level of metalinguistic reporting. However, while this distinction is clearly significant, perceptual measures of sociolinguistic knowledge *below* the metalinguistic level can themselves vary a great deal. Campbell-Kibler (2012) included a matched guise social evaluation task as an intermediary between a more explicit measure (self-report) and a more implicit measure (implicit association task), finding that individuals' responses in the matched guise task correlated with neither the metalinguistic nor the implicit association task measure. This suggests that differences among *implicit* tasks can differently access sociolinguistic knowledge, which raises the question of what aspects of the tasks may be responsible for these differences.

In the realm of social cognition, Devine (1989) and others (e.g., Lieberman, 2007; Schneider & Shiffrin, 1977) noted that processing time mediates whether a perceiver has enough control to allow for personal beliefs and desires to intervene and alter or inhibit automatic responses (Devine, 1989:6). Paradigms commonly used to examine *implicit* sociolinguistic knowledge alone can vary dramatically with respect to the amount of processing time afforded to a participant, as well as the nature of the task at hand. However, little is known about whether different "implicit" paradigms access sociolinguistic knowledge in the same way. For example, though the linguistic manipulation remains implicit in the matched-guise social evaluation tasks commonly used in work on sociolinguistic perception (e.g., Campbell-Kibler, 2007; Giles, 1970; Lambert, Hodgson, Gardner, & Fillenbaum, 1960), participants are given a relatively extended amount of time and degree of conscious control over their responses. In paradigms that assess the influence of social information on a linguistic task such as phoneme categorization (e.g., Hay et al., 2005), or sentence completion (e.g., Squires, 2013; Staum Casasanto, 2008), listeners typically provide a more rapid response to each stimulus, though the choice is made consciously. Finally, earlier and more automatic levels of sociolinguistic perception are accessible in measures newer to sociolinguistic perception, such as the implicit association task (Campbell-Kibler, 2012) or eye-tracking (Koops et al., 2008). Examining how timing and task may influence implicit responses can help reveal details of the processes of sociolinguistic perception entirely below the level of conscious awareness. This can help elucidate the ways in which social and linguistic information are linked in perception and processed in relation to one another (e.g., Sumner, Kim, King, & McGowan, 2014).

This paper begins to address this issue by directly examining how implicit knowledge of a feature-meaning link, itself not subject to conscious commentary

and thus classifiable as a *marker*, arises in measures that access different stages in the time course of processing and that allow for different degrees of control over responses. To draw the distinction between explicit and implicit sociolinguistic knowledge, the studies previously mentioned tested features that at least some participants can discuss metalinguistically—*stereotypes*, in Labov's (1972) tripartite distinction. This adds a potential confound to the comparison of different implicit tasks, such that explicit, metalinguistic knowledge might interfere with some implicit measures more than with others. The comparison of more controlled versus more automatic perceptions of a *marker*, entirely below the level of metalinguistic commentary, provides a step toward understanding how responses at different points in processing time and at levels of automaticity operate *within* the domain of implicit sociolinguistic perception, when metalinguistic awareness of the variable is not available.

In this study, I deploy two contrasting experimental paradigms to test implicit listener awareness of a link between one linguistic feature (backing of the TRAP vowel in American English) and its social associations with Californian origin, the valley girl persona, and the business professional persona. Through a comparison of three types of responses to the same auditory stimuli of TRAP-backing, I assess the relation between slower and more deliberate versus faster and more automatic processes of sociolinguistic perception that occur in response to a sociolinguistic variable about which there is little to no explicit commentary. The two experiments used in this study—a matched guise social evaluation paradigm versus a forced lexical choice paradigm using eye tracking—were selected because they measure three different degrees of response automaticity, but none require conscious or metalinguistic acknowledgment of the feature-meaning link. By juxtaposing results from social evaluations, lexical categorizations, and early eye movements, I demonstrate that a *marker* may be activated for a listener differently depending on the time point in processing, the type of task, and the level of deliberative control over responses that a listener is able to bring to that task. Findings from these different measures do not neatly mirror one another, suggesting that different points in the time course of processing and different response types allowed by a task may tap into different aspects of sociolinguistic knowledge. Based on these results, I suggest that conclusions about implicit sociolinguistic knowledge, and the behavioral results of this knowledge, must be made in light of the response type at hand.

TRAP-BACKING

The macrosocial associations of TRAP-backing have been established via work on regional dialects, which has shown that TRAP is backing as part of the California Vowel Shift among speakers across the state of California (D'Onofrio, Eckert, Podesva, Pratt, & Van Hofwegen, 2016; Hinton, Moonwomon, Bremner, Luthin, Van Clay, Lerner, & Corcoran, 1987; Kennedy & Grama, 2012) and other areas of the Western dialect region (e.g., Becker, Aden, Best, & Jacobson,

2016). However, unlike the metalinguistic commentary that arises around pronunciation of this phoneme in locales like the U.S. Inland North (e.g., Driscoll & Lape, 2015), explicit commentary on this feature from Californians is virtually nonexistent. In interviews throughout the Central Valley of California (e.g., D'Onofrio et al., 2016), no individuals explicitly recognized or described the feature to the author, even when prompted. TRAP-backing has also been deployed in depictions of the valley girl persona—a young, female persona named for Los Angeles county's San Fernando Valley who is defined as shallow, materialistic, and unintelligent. The valley girl and her associated linguistic style—Valspeak (Donald, Kikusawa, Gaul, & Holton, 2004)—has been popularized through parodies such as Frank and Moon Unit Zappa's 1982 song, "Valley Girl," and Saturday Night Live's 2012–2013 skits, *The Californians*. While internet searches turn up no discussion of the TRAP vowel in metalinguistic discussion of Valspeak, TRAP-backing *has* been found in parodic performances of this persona (Hinton et al., 1987; Pratt & D'Onofrio, 2017). Note that this persona is largely stereotypical in nature—individuals do not tend to self-identify as valley girls, as it is typically used in a derogatory fashion. The parodic depictions may reflect some features used by young women in the San Fernando Valley area, but these are almost certainly exaggerated in these parodic performances.

In contrast to the valley girl persona, TRAP-backing may also serve as an index of educated, formal, or professional speech, perhaps by virtue of supraregional movement away from stigmatized TRAP-raising (e.g., Driscoll & Lape, 2015) or via an indexical association with British English. Podesva, Hall-Lew, Brenier, Starr, and Lewis's (2012) study of Condoleezza Rice demonstrated the potential for TRAP-backing to index a formal, educated, "correct" way of speaking, showing a greater degree of lowering and backing of preobstruent TRAP in Rice's formal scripted speech context as compared to the less formal question-and-answer context. These persona-based associations with TRAP-backing have also been demonstrated in listener perceptions (D'Onofrio, 2015a; Villarreal, 2016), which show that listeners associate the valley girl and/or the business professional with a backer production of TRAP. The present paper tests these same implicit feature-meaning links through different types of responses: social evaluations, lexical categorization of a linguistic stimulus, and early eye movements. This allows for an investigation of whether the same associations arise at these different time stages in processing and also allows for a comparison between these three social meanings across tasks.

EXPERIMENT 1: MATCHED GUISE TASK

Experiment 1 deploys a matched guise technique to assess how listeners socially evaluate TRAP-backing, asking specifically whether the feature corresponds to California-ness, a valley girl persona, and/or a business professional persona. In a matched guise task that examined TRAP-backing and GOOSE-fronting, both features of the California Vowel Shift, Villarreal (2016) found that guises with

both features corresponded to higher ratings of Californian origin and valley girl–ness. In experiment 1, I isolate TRAP-backing to assess its influence alone. The matched guise technique elicits listener evaluations of the same speaker using sets of stimuli that vary in some linguistic feature or set of features (Lambert et al., 1960). Evaluations allow for relatively slower, offline responses about social associations with a voice, as compared to experiment 2, which accesses perceptual behavior that is quicker and more automatic, and can be assessed at an earlier stage in processing.

Stimuli

The critical auditory stimuli used in both experiments consist of individual words resynthesized from read productions of TRAP-LOT minimal pairs. Continua from TRAP to LOT words were constructed by manipulating read utterances produced by the author, a native speaker of American English in her mid-20s at the time of recording, from the Northern dialect region[1] (Labov, Ash, & Boberg, 2006). The speaker was recorded in a soundproof booth using a Turner 2302 microphone with a Rane MS1b preamplifier. Recordings were digitized (44.1 kHz, 24 bits) with an Edirol UA-101, recorded into the software program Audacity. The speaker read a list of monosyllabic (consonant)consonant-vowel-consonant TRAP-LOT minimal pairs.

Nine-step continua from each of the recorded TRAP tokens to respective LOT tokens were then created using the Akustyk package (Plichta, 2013) in Praat (Boersma & Weenink, 2011), with the command *Create speech continuum*. This command takes two vowels as input and creates tokens at the mean duration of the two original tokens while resynthesizing the original TRAP's F1, F2, and F3 values to progressively match the values of the original LOT token in nine equal steps. The pitch of all resultant tokens and formant values above F3 matched the original TRAP token. All vowel tokens were resampled to 10 kHz, 16 bits—the preceding and following phonological frame and all filler tokens were likewise resampled. Manipulated tokens were embedded in the preceding and following frames from the original TRAP token, though for preceding fricatives (including aspiration in voiceless stops), a nine-step fricative continuum was created by Praat script from TRAP to LOT words' fricatives and matched to the corresponding vowel continuum step. All steps in the continua were scaled for peak amplitude. Naturalness ratings of each of these words were collected from 10 online participants per word, on a sliding scale from 1 "sounds like a human" to 10 "sounds like a computer, or manipulated." All recordings used were given an average rating of less than 3. While these ratings varied by item, there was no consistent pattern in these ratings by backness. For the matched guise technique, two points on the resynthesized continua were selected as stimuli: the middle point on each continuum, ambiguously interpretable as a backed TRAP or as a fronted LOT, and the frontest point on each continuum. This yielded a set of eight words with two backness variants each (Table 1).

TABLE 1. *Manipulated formant measurements for resynthesized critical stimuli*

Word	Backed token			Fronted token		
	F1	F2	F3	F1	F2	F3
BLACK	968	1649	2983	1057	1950	3013
HAT	1162	1832	3026	1021	2177	3148
LACK	1123	1838	2918	1117	2043	2889
PAT	993	1853	2932	997	2041	2983
RACK	1044	1887	2844	1117	2062	2896
SACK	1005	1603	2718	1162	1966	2746
STACK	980	1622	2552	983	1877	2571
TAP	930	1627	2687	997	1861	2795

Filler words were also recorded in the same session. These words had the same phonological structure as the critical stimuli and contained either a FLEECE, KIT, FACE, or DRESS vowel. Two word lists were created for experiment 1 (Table 2). The first contained all eight of the critical TRAP stimuli interspersed with the eight filler words, aiming to mask the variable of interest. The second list contained only the eight TRAP stimuli, which drew attention to the particular vowel of interest, but eliminated the possible influence of the filler productions on interpretations of the feature of interest. For both lists, two separate auditory samples were created such that listeners heard either all front or all back tokens of TRAP.

Four total auditory samples were used in the matched guise task, assigned between subjects: (a) backed TRAP tokens with fillers, (b) backed TRAP tokens with no fillers, (c) front TRAP tokens with fillers, or (d) front TRAP tokens with no fillers. All four auditory samples were created through a concatenation of the words in Praat with 500 msec of silence between each word. Note that since all participants heard the entire list of words in the same order as one auditory sample, then provided one social evaluation of the voice following the entire sample, it is not possible in this particular task to disambiguate whether any particular TRAP word may have been responsible for the effect or whether it was a cumulative effect.

Procedure

Each participant was presented with one of the four auditory samples. Listeners were recruited and compensated online via Amazon's Mechanical Turk, a crowd-sourcing web service (Schnoebelen & Kuperman, 2010). Participants were directed to a Qualtrics survey through which they first read a consent form, then completed an audio check. Listeners were then presented with a page containing instructions, the auditory sample, and ratings. Listeners were told, "We are interested in what impressions you can get about a person just based on their voice. You will listen to a person reading a list of words. This person is from the United States and is an American English speaker." They were then asked to click the play button to hear the list. An orthographic transcription of

TABLE 2. *Matched guise task auditory word lists and total time of auditory samples*

Word list with fillers	*fate, slip, black, chick, hat, lack, pat, chip, creek, feet, rack, sack, flit, sleep stack, sit, tap, steep* (22 sec)
Word list with no fillers	*black, hat, lack, pat, rack, sack, stack, tap* (7 sec)

the word list was presented alongside the auditory sample, and listeners were instructed to follow along with the list as they were reading, to ensure that they identified each word, particularly critical for backed tokens of TRAP that could be confusable with a token of LOT. Participants were asked to listen to the *entire* list before answering any questions, and they could play the sample as many times as they liked. To ensure that participants listened to the clip, they were asked to type as many words as they could remember following the main rating task. Only those who reported at least one TRAP word accurately were included in the analysis.

Listeners were then asked to rate the speaker's demographic characteristics and characterization as a set of social types. For location of origin, listeners were asked to select from Alabama, California, Illinois, New York, or Oklahoma. Note that the only state in this set that exhibits state-wide TRAP-backing is California, also the only option within the Western dialect region. Personae were included in the evaluations to test the valley girl and business professional meanings of TRAP-backing. Listeners rated the speaker on a five-point likelihood scale that the speaker was that type of person, from very unlikely to very likely. After completing their evaluations, listeners reported the listening device used and the surrounding noise level. They then completed a demographic questionnaire, in which they were asked to self-identify their age, gender, native language(s), locations lived, and the ages at which they lived there.

Participants

Listeners received $.80 for completing the task, which took an average of 5.3 min to complete.[2] After eliminations, data from a total of 196 participants was analyzed (constituting 196 responses per question). Listeners were categorized into Western versus non-Western dialect regions (Labov et al., 2006). Participant age, gender, and dialect region by auditory stimulus in experiment 1 are provided in Table 3.

Results

I analyze the influence of TRAP backness on perceived speaker origin (California versus another state) and likelihood ratings of the speaker as personae previously associated with TRAP-backing—valley girl and business professional. Scalar ratings for persona likelihood were normalized by listener to control for potential listener-based differences in range, by calculating *z*-scores for the likelihood scales based on all the likelihood ratings provided, including those for filler personae, leaving ratings centered on zero. Results for perceived location of

TABLE 3. *Participant background information, by condition*

	Backed no fillers	Backed fillers	Fronted no fillers	Fronted fillers
Mean age	38	37	34	38
Gender				
Female	18	13	24	28
Male	32	32	23	24
Another	0	1	0	1
Western	16	13	14	13
Total	50	46	47	53

origin were analyzed by fitting a logistic regression model[3] on the binary dependent variable of California selection versus non-California selection (any of the other four states). Backness of TRAP (back versus front) and list type (fillers versus no fillers) were included as fixed effects. The fixed effects that served as manipulations in the experiment (backness, fillers) were retained in all models, with fillers included to assess effects of backness regardless of whether fillers were included. Participant factors were tested as fixed effects in all models. For all of the following models presented in this paper, participant factors were not included in the final model where they did not improve model fit.[4] Summaries of the fixed effects of the simplest best-fit models are provided. The logistic regression model predicting selection of non-Californian origin is shown in Table 4.

While the speaker was generally heard as a Californian at rates above chance (20%) for all auditory samples (no fillers, back = 32%; no fillers, front = 30%; fillers, back = 37%; fillers, front = 40%), this rating was not modulated by backness of the TRAP tokens (Figure 1). The model did show listener origin to be a significant predictor of Californian ratings for this voice, indicating that listeners from the Western dialect region were more likely to rate the speaker as Californian than non-Western listeners were, which perhaps reflects a tendency of Westerners to be more likely to select the Western state in general. However, dialect region did not interact significantly with TRAP-backness, indicating that regardless of the origin of the listener, backness did not have an effect on regional ratings of the speaker.

The likelihood that this speaker was a valley girl was assessed using a linear regression model predicting the normalized likelihood rating with fixed effects of TRAP backness and list type (Table 5). Participant information predictors did not improve model fit and were thus not included in the model. Results show that regardless of fillers, participants who heard backed TRAP tokens provided a higher mean rating for likelihood that the speaker was a valley girl than participants who heard front TRAP.

Furthermore, in the model presented in Table 5, I include as a binary predictor whether or not the participant selected California for the speaker's state of origin. This predictor was significant, indicating that listeners who thought the speaker was

TABLE 4. *Non-California speaker origin selection, logistic regression summary of fixed effects* (n = 196 responses)

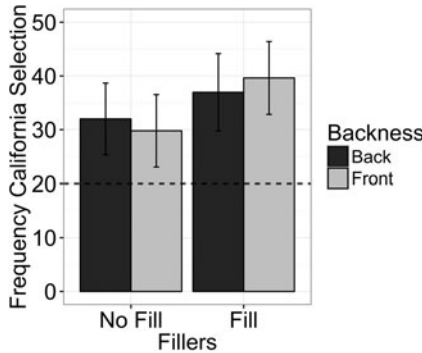| Predictor	Estimate	SE	z-Value	p-Value	% Non-California selection
(Intercept)	1.12	.30	3.73	<.001***	
Backness (default = back)					66
Backness = front	−.04	.31	−.13	.901	65
List type (default = no fillers)					69
List type = fillers	−.38	.31	−1.24	.216	62
Listener Western origin (default = non-West)					71
Listener Western origin = West	−.841	.33	−2.56	.011*	52

Note: *p < .05; ***p < .001.

FIGURE 1. Frequency California selected for speaker's state of origin, by backness of TRAP tokens and filler condition. Dotted line indicates chance selection.

TABLE 5. *Normalized likelihood rating for valley girl persona, linear regression summary of fixed effects* (n = 196 responses)

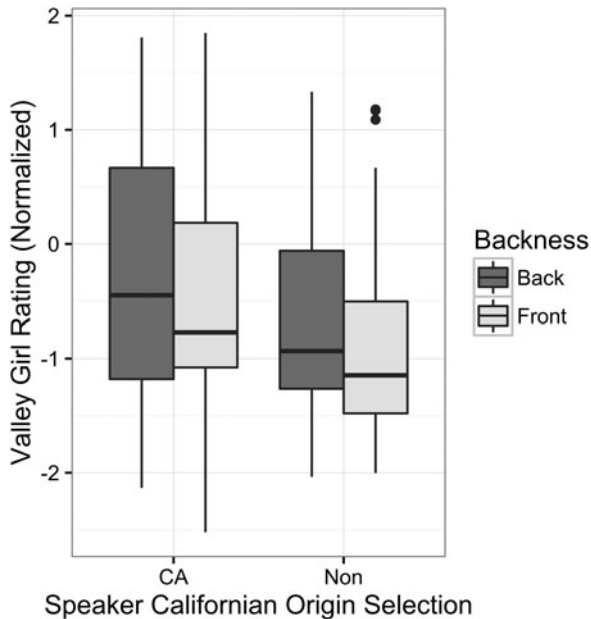| Predictor	Estimate	SE	t-Value	p-Value	Mean valley girl rating (normalized)
(Intercept)	−.25	.15	−1.69	.093	
Backness (default = back)					−.46
Backness = front	−.31	.13	−2.38	.018*	−.75
List type (default = no fillers)					−.71
List type = fillers	.19	.13	1.41	.162	−.51
Perceived speaker origin (default = California)					−.31
Perceived speaker origin = non-California	−.45	.14	−3.25	.001**	−.77

Note: *p < .05, **p < .01.

FIGURE 2. Normalized likelihood rating for valley girl persona, by backness of TRAP tokens and speaker state of origin selection.

from California gave the speaker a higher likelihood rating of being a valley girl than those who did not think the speaker was from California, regardless of the stimulus heard (Figure 2).

Though it is not possible to say the directionality in which this operated, there is a very clear association between California and the valley girl persona across the board, and the relation between TRAP backness and the valley girl rating persists for both listeners who rated the speaker as a Californian and those who did not (Figure 2). Although the two social factors are highly correlated in participants' evaluations, TRAP backness itself modulated ratings for only the valley girl persona, not for broader U.S. state selection.

Finally, I examine ratings of the likelihood that the speaker was a business professional. A linear model was fit to examine the association between backness and the normalized business professional ratings using the same methods, predictors, and method of model comparison used for the valley girl ratings (Table 6). No statistically significant effect of backness emerged, suggesting that these backed tokens are associable with the valley girl persona and not with the business professional persona.

Overall, results from experiment 1 most clearly show TRAP-backing's association with the valley girl persona. However, even though this persona was clearly correlated with Californian origin for these listeners, the macrosocial association between California as a state and TRAP-backing was not directly activated in

TABLE 6. *Normalized likelihood rating for business professional persona, linear regression summary of fixed effects (*n = 196 responses)

Predictor	Estimate	SE	*t*-Value	*p*-Value	Mean business professional rating (normalized)
(Intercept)	−.01	.12	−.11	.910	
Backness (default = back)					−.07
Backness = front	.24	.13	1.75	.081	.16
List type (default = no fillers)					.10
List type = fillers	−.12	.13	−.91	.362	−.01

social evaluations, even for listeners who were themselves from the Western dialect region. While ideologies related to the valley girl persona have allowed for a dislocation of this persona to characterize a younger female of various regions of origin, she is most notably associated with Southern California—indeed, she is named for Los Angeles county's San Fernando Valley. The significant relation between the valley girl rating and the selection of California origin demonstrates this ideological link between the two, as also shown in Villarreal (2016), though the present results illustrate that TRAP-backing alone may not cue general California-ness in social evaluations.

In experiment 2, I turn to two measures that assess the influence of these same social associations with TRAP-backing on (a) lexical categorization of ambiguous stimuli between TRAP and LOT and (b) a very early perceptual response to these stimuli prior to categorization, as measured by eye movements. The links between social information and TRAP-backing that are found in social evaluations (experiment 1) do not neatly predict the links that will arise in early and automatic responses (experiment 2), suggesting that these measures, while both assessing implicit sociolinguistic knowledge, do not mirror one another. Instead, different social associations of TRAP-backing can be foregrounded in responses at different time stages of processing and in different tasks.

EXPERIMENT 2: EYE MOVEMENTS

Experiment 2[5] uses a four-alternative forced choice categorization task with eye-tracking. In this paradigm, eye gaze serves as a proxy for the time course of the listener's decision-making process: the points at which a listener's eyes are focused on a screen indicate the options they are considering when categorizing a speech signal as a particular lexical item. The use of eye-tracking has risen to prominence in work on sentence processing (for an overview, see Huettig, Rommers, & Meyer, 2011; Rayner, 1998; and Tanenhaus, 2007) as a measure of participants' shifting expectations across the entire time course of online linguistic perception. While participants have some degree of control over where they fixate their eyes, gaze is more automatic than a deliberate mouse click or

key press. Most importantly, it can be used to measure reactions at extremely early time points in processing—eye gaze fixations can reflect exposure to a stimulus as early as 200 msec from the onset of that stimulus (Allopenna, Magnuson, & Tanenhaus, 1998). In experiment 2, analysis examines the influence of social information on responses immediately following a listener's hearing of a linguistic stimulus and on how listeners ultimately categorize that stimulus, through a mouse click.

Auditory stimuli

Critical auditory stimuli for experiment 2 were taken from those used in experiment 1. Since the comparison in this paradigm is among listener responses to the same linguistic stimuli given different social primes, this study included only the backed tokens of TRAP from experiment 1. Here, a comparison to the fronted stimuli was not included, as no social prime differences were expected, though it is an area for future work. The filler stimuli described in experiment 1 were also used as fillers in experiment 2. In addition, filler stimuli were recorded by a second, male voice to serve as a distracter to the variable of interest. The male speaker, an American English speaker in his 30s from the Inland North region of the United States, recorded a list of filler words in the same manner as described in experiment 1. All filler words were scaled for peak amplitude and resampled to match the resynthesized auditory stimuli.

Procedure and design

The design of this task used a "visual world" of four words on a screen (Dahan, Drucker, & Scarborough, 2008; Huettig et al., 2011; Koops et al., 2008), in which listeners saw four orthographic words, heard an auditory word, and clicked on the word they heard in each trial. The experiment was presented through EPrime, using a ToBII T606XL remote eye tracker. Participants were seated in front of a monitor, and the task began with a calibration of the eye tracker to the participant's eyes and eye movements. After a successful calibration, the following instructions were provided on the screen: "In this experiment you will see four words. Please examine these words until a picture appears in the center of the screen. The picture will represent the person you will hear speaking." Participants were then presented with instructions that they would hear two voices, and they were shown visual icons corresponding to the voices.

Listeners were placed in one of four social information conditions: one group was given no speaker information (baseline condition), a second group was told that the speaker was from California, a third group was told the speaker had been described as a valley girl, and a fourth group was told the speaker had been described as a business professional. Listeners were shown one of the icons in the first row of Figure 3 to represent the critical voice, corresponding to the social prime condition they were assigned. The corresponding icon in the second row of Figure 3 served as an icon representing the male distracter voice.

Baseline   Region   Persona

Critical
Voice

Green circle   'From   'Described as   'Described as
(no written info) California'   a Valley Girl'   a Business
                                                   Professional'

Distracter
Voice

Orange circle   'From   'Described as a
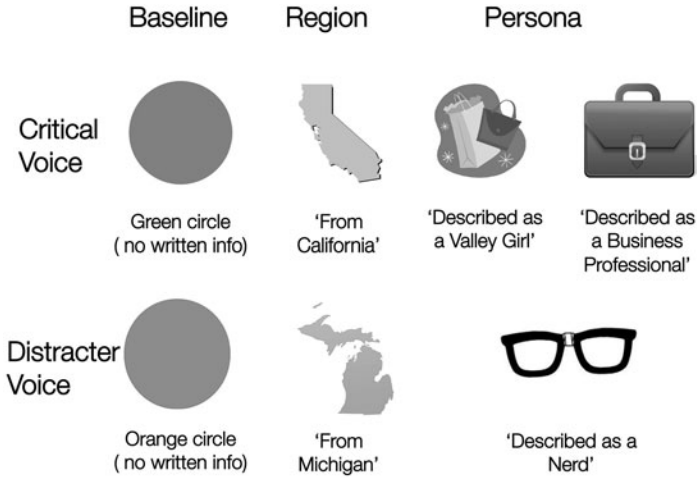(no written info) Michigan'   Nerd'

FIGURE 3. Icons used to correspond to social information conditions.

Listeners were told social information about the speakers explicitly in a written instruction (e.g., "One speaker has been described as a valley girl and will be represented by the following picture [shopping bag icon]; the other speaker has been described as a nerd and will be represented by the following picture [glasses icon].") While these icons are included to remind the listener of the social prime, the picture alone was not necessarily intended to evoke the social meaning without this written prompt. Analyses compare listeners in the baseline condition with those in each social prime condition. Listeners were then told: "When you focus your eyes on the picture, you will hear a word. Listen carefully and look at the word you hear. Click on that word as quickly as you can." Participants then completed four practice trials.

In each trial, four words were presented on the screen, one in each corner. The four words composed two sets of minimal pairs, each pair beginning with a different onset (e.g., *sack–sock*, *leak–lake*). One of the orthographic words was a target word, or the word that matched the auditory stimulus (e.g., *sack*). Another was a competitor word, or a word that was different from the auditory stimulus in vowel sound only (e.g., *sock*). The remaining two words were members of a distracter pair (e.g., *leak* and *lake*). In each trial, all words contained the same number of letters, and each minimal pair differed phonologically only in the vowel. For each trial, participants were first presented with the words and allowed to familiarize themselves for 5 sec. Then, the icon corresponding to the upcoming voice appeared in the center of the screen. Once participants fixed their gaze on the icon, an auditory stimulus was played. Listeners then used the mouse to click on the word that they heard, advancing them to the next trial.

Of the 32 total trials presented to each listener, four were critical trials, in which the auditory token was a backed TRAP token from experiment 1, and the corresponding orthographic TRAP word was posed against a LOT competitor.

These were the trials of interest for the present analysis—specifically whether or not listeners clicked on or looked to the TRAP word (e.g., *sack*), an indication that they were hearing the token as an instance of backed TRAP (as opposed to a token of LOT). An additional 12 filler trials were presented in the same voice as the critical trials, including both the non-TRAP filler auditory stimuli used in experiment 1, and trials where the auditory stimulus was a backed TRAP token but the selection was not a critical contrast (e.g., *sack–sick*). The remaining half of the trials (16) consisted of responses to tokens produced by the male distracter voice. Distracter voice tokens did not include TRAP or LOT vowels, and every participant completed the same 16 distracter voice trials.

Two trial lists were created to counterbalance the pairing of auditory stimulus and trial screen across participants. Each participant heard each auditory stimulus once, and placement of each word on the screen was balanced across trials such that the target word appeared in each of the four corners of the screen at the same rate. Minimal pairs always appeared adjacent to one another, either side by side, or vertically (never diagonally from one another). The order in which the trials were presented was randomized for each participant. Only critical trials were analyzed.

Participants

For this study, 39 participants, all self-reported native speakers of American English, were recruited via Stanford University subject pools and were either compensated with course credit or with $7, and the experiment typically took 15–20 min. After the main experiment, participants completed a demographic questionnaire. Locations lived were coded according to dialect region and as a binary West versus non-West variable, as in experiment 1.[6] Participant information by social information condition is provided in Table 7. Again, participant factors were tested as fixed effects in all of the following regression models and were removed only where they did not improve model fit.

Word choice

The main task for listeners in experiment 2 was to categorize ambiguous stimuli as either a TRAP or LOT word. Selection in critical trials, as measured by mouse click, took an average of 1.37 sec. Word choice between TRAP and LOT was analyzed in critical trials. Five trials were removed because response times were above 2 SD of the mean response time or selection of a distracter word was made. This measure assesses faster and less controlled responses to the stimuli than those in in experiment 1, but less automatic than the early eye-gaze measure to follow. Word choice in these trials was analyzed statistically via a mixed-effects model with social information condition and trial order as fixed effects, random intercepts for participant and item, and random slopes for trial number by participant and item. The final best-fit model is shown in Table 8.

No significant differences emerged between the four conditions in terms of word choice. Listeners who were told that the speaker had been described as a

TABLE 7. *Participant background information, by condition*

	Baseline	California	Valley girl	Business professional
Mean age	20	21	20	25
Gender				
Female	5	7	5	5
Male	5	2	5	5
Western	6	5	6	7
Total	10	9	10	10

TABLE 8. *Mixed-effects regression summary of fixed effects for choice of* TRAP *(versus* LOT*) word (n = 151 responses)*

Predictor	Estimate	SE	*t*-Value	*p*-Value	Percentage of TRAP selections
(Intercept)	.70	.11	6.49	<.001***	
Condition (default = baseline)					75
Condition = business professional	−.01	.08	−.07	.942	72
Condition = California	.14	.09	1.66	.101	89
Condition = valley girl	.15	.08	1.91	.062	93
Western dialect region (default = non-West)					88
Western dialect region = West	−.14	.06	−2.22	.031*	78
Trial number	.01	.00	2.18	.036*	

Note: *$p < .05$, ***$p < .001$.

valley girl were most likely to categorize a given ambiguous token as TRAP, more frequently than those in the baseline condition in which listeners were given no social information. Listeners in the California condition also heard ambiguous tokens as TRAP more frequently than those in the baseline condition, as expected. The business professional condition did not show a difference from baseline in word choice.

Trial number significantly predicted word choice, such that later trials showed a greater likelihood of TRAP selection. An interaction between trial number and social prime condition was not significant, indicating that this tendency to choose TRAP more over the course of the experiment persisted regardless of the social prime. Participant Western background significantly affected responses in a direction that was unexpected based on previous literature and experiment 1. Here, listeners who were from the West were less likely to select the TRAP word than those who were not, perhaps a result of the differing participant populations in the two experiments. While non-Western listeners in experiment 1 had not lived in the West at all, non-Western listeners in experiment 2 all lived in California at the time of test, and the association between TRAP backness may have been heightened in salience for these speakers, transplants to a TRAP-backing environment.

While more robust data is required to assess whether the trend of social prime is borne out, effects of social prime condition in some ways correspond to the results of experiment 1. The valley girl prime led to the greatest increase in expectations of TRAP-backing (18 percentage points higher than baseline), suggesting that in categorizing an ambiguous speech stimulus, listeners show an association between the valley girl persona and TRAP-backing, a link observed in experiment 1. The California social prime also trended in this direction (14 percentage points higher than baseline). Thus, we see indications of these associations in a task that primes a listener with social information and requires a lexical choice. In the next section, I examine listeners' responses to these linguistic stimuli in the earliest window following exposure to the critical auditory stimulus in this same task.

Early eye movement results

In order to analyze earlier and less controlled responses to these stimuli, I focus here on where on the screen listeners look in a time window immediately following presentation of the ambiguous vowel. This assesses how primed social information may modulate the earliest and most automatic reactions that listeners have to a speech signal, prior to a deliberate lexical categorization. If social information leads listeners to *automatically* expect TRAP-backing, we would expect listeners with that social information to be more likely than listeners with no information to look to the target TRAP word more rapidly upon hearing the ambiguous vowel.

Gaze was measured from onset of the auditory stimulus to the point of decision (the mouse click). The eye tracker recorded fixations on the screen as *x-y* coordinates throughout each trial for both eyes, with analysis conducted on participants' left eye fixations. Coordinates were then coded categorically to correspond to which of the four words on the screen was being fixated. Fixations to a word were only counted if the word itself was being fixated. Saccades (movements between fixations) were not analyzed. Fixations in critical trials were coded as a fixation to the target word (TRAP word), the competitor word (LOT word), distracter (either of the words in the distracter pair), or no word. Given that ultimate lexical choice differed across the four conditions, I assess here data from only trials where TRAP was selected—those in which listeners ultimately decided that they were hearing a backed TRAP. Here, I measure fixations to the target (TRAP) word within the earliest window in which a listener reaction to the vowel itself was possible. Accounting for the approximately 200 msec required for saccade planning and execution (Allopenna et al., 1998), this window began at 200 msec following onset of the vowel in the auditory token, and ended at 200 msec following the offset of the vowel (between the dashed lines shown in Figure 4). Analysis of results presented in Table 9 and Figure 4 assess only this early time stage, occurring an average of 1000 msec prior to the mouse click assessed.
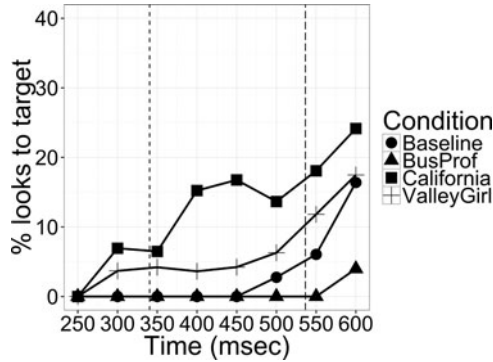
FIGURE 4. Early time course of proportion of looks to the target (TRAP) by social information condition, averaged across TRAP-selection trials only. Time from onset of auditory word. Dashed lines indicate time window analyzed (mean vowel onset + 200 msec to mean vowel offset + 200 msec).

TABLE 9. *Mixed-effects regression summary of fixed effects predicting proportion looks to target (TRAP word) in time window between 200 msec following vowel onset and 200 msec following vowel offset, TRAP-selection trials only (n = 122 trials)*

Predictor	Estimate	SE	*t*-Value	*p*-Value	Percentage of looks to target
(Intercept)	−.04	.04	−.84	.407	
Condition (vs. baseline)					3.7
Condition = business professional	−.01	.05	−.10	.925	1.8
Condition = California	.18	.05	3.42	.001**	11.9
Condition = valley girl	.06	.05	1.29	.205	5.8
Participant gender (vs. male)					5.1
Participant gender = female	.07	.04	1.91	.067	6.6

Note: **$p < .01$.

Within this time frame, a time window analysis was conducted (following Dahan et al., 2008; and Koops et al., 2008).[7] Proportion fixations to the target (TRAP) word within this early time window were calculated by trial and used as the dependent measure in a linear mixed-effects model (Table 9). A fixed effect of social information condition was included (default = baseline condition), with random effects of participant and item. Trial order was tested in the model, but had no effect on early eye movements and its inclusion did not improve model fit. While no participant factors yielded significant results, nor did they improve model fit, participant gender did have a marginal effect on eye movements and marginally improved the model. Given that participant gender is not evenly balanced across the conditions (Table 7), I retained participant gender in the model to examine social prime effects that emerge controlling for gender.

A significant difference emerged between the California condition and the baseline condition in the immediate gaze responses after the vowel was played. This indicates that even for listeners who all ultimately made the same word selection, those who thought that the speaker was a Californian were significantly more likely to look to the TRAP word immediately than those who had no information. This social information thus modulates very early reactions to a linguistic stimulus. No significant differences emerged between the baseline condition and the valley girl nor business professional conditions, though the valley girl condition trended in the expected direction to a small degree. A post hoc comparison among the social primes was conducted via Tukey's honestly significant difference test on the condition factor of the linear model, with a Bonferroni-Holm correction. The California and valley girl conditions were marginally significantly different from one another according to this comparison (est. = 0.12; $z = -2.31$; $p = .082$).

The early eye-tracking measure reveals listeners' associations between TRAP-backing and California, with the association between the valley girl persona and TRAP-backing attenuated compared to the effect of the California prime at this early stage. It ultimately emerges in word choice in this same task, however, overtaking the effect of California information. As in experiment 1, there is no evidence here of a link between the business professional persona and an expectation of TRAP-backing in either measure.

DISCUSSION AND CONCLUSIONS

The experiments presented in this paper examine listeners' social associations with TRAP-backing, a *marker*, using three measures of implicit sociolinguistic knowledge that access different amounts of processing time and degrees of listener control over responses. Results demonstrate that in social evaluations and in more controlled and more automatic measures of socially primed linguistic perceptions, listeners associate TRAP-backing with California-related social meanings, either with Californian origin broadly construed, or with a Californian persona, the valley girl. The link between TRAP-backing and the valley girl persona arose in the slowest and most introspective measure—social evaluations of linguistic stimuli used in experiment 1, in which listeners who heard words with backed TRAP rated the speaker as more likely to be a valley girl than listeners who heard more fronted TRAP vowels. Though ratings of Californian origin predicted significantly higher valley girl ratings, there was no association between perceived California origin and TRAP-backing in evaluations, even among those who were from a Western TRAP-backing dialect region. An apparent link between the valley girl persona and TRAP-backing also arose in experiment 2, in the patterns for word choice: the expectation of a valley girl persona makes listeners more likely to classify a token ambiguous between TRAP and LOT as TRAP (93% as compared to 75% at baseline). Finally, in early eye movements tracked in the *same task*, listeners who thought the speaker was a Californian were significantly more

likely than listeners with no speaker information to look toward the TRAP word upon first hearing an ambiguous TRAP-LOT stimulus, indicating an early and relatively automatic expectation of TRAP-backing from a Californian speaker.

Prior work has illustrated mismatches between listeners' metalinguistic and implicit sociolinguistic knowledge (Campbell-Kibler, 2012; Levon & Fox, 2014; Squires, 2016), supporting models of sociolinguistic perception by which listeners' sociolinguistic representations at metalinguistic or conscious levels may differ from associations stored at more implicit levels (e.g., Campbell-Kibler, 2012). By focusing on a sociolinguistic *marker*, the present study probes how time applied to processing a linguistic input and the type of response elicited may modulate how implicit sociolinguistic knowledge is demonstrated, without any potential influence from explicit, metalinguistic knowledge of a feature. In each of the measures tested in this study, evidence emerges that listeners implicitly link TRAP-backing with California-ness. However, the type of social information that is linked with a backed TRAP differs across the three measures. Results suggest that an early, automatic response to a socially meaningful linguistic stimulus may be shifted, attenuated, or perhaps elaborated after a listener has more time with which to process the stimulus and more ability to control their responses. And crucially, this shift does not require metalinguistic knowledge of the sociolinguistic feature being perceived. In experiment 2, the macrosocial California prime shows a significantly greater influence than the valley girl or baseline primes in responses at the earliest and most automatic level tested. However, the persona-based valley girl prime marginally influenced the later stage of lexical choice in this same task, as measured by a mouse click that occurred about 1 sec later, on average. This suggests that listeners maintain implicit representations linking a linguistic form and a social meaning that can be expressed differently given different amounts of time and control to apply to linguistic processing. Further supporting this notion, the valley girl persona information was significantly tied to TRAP-backing in more introspective social evaluations (experiment 1), the measure that allowed the greatest amount of time and control of the three, though macrosocial California origin showed no relation to TRAP-backing.

One limitation in the comparison between experiments 1 and 2 is the different participant groups: in experiment 1, participants were recruited from throughout the United States, while in experiment 2, all participants were residing in California at the time of the experiment. However, two factors suggest that the patterns observed between these tasks are related to the response type, rather than simply to participant-based differences. First, while experiment 1 included participants from a variety of dialect regions, and region influenced social evaluation responses overall, participant region of origin did not interact with TRAP backness in either valley girl ratings nor in speaker's imputed California origin. That is, even listeners in experiment 1 who were themselves from a TRAP-backing region (including California) did not show a correlation between backed manifestations of TRAP and Californian state selection, nor did they show different behavior from non-Western participants in their positive linkage

between TRAP-backing and valley girl likelihood ratings. Second, the two different measures within experiment 2—responses from the same participants, in the same task—differed in the degree of influence that the California prime versus the valley girl prime illustrated. That is, the social prime that most strongly influenced listeners' early reactions was not the same as the social prime that most strongly influenced those same listeners' ultimate word choice. Here, only time and response type can be responsible for the differential effects between the two social primes.

The contrast in responses across the measures indicates that paradigms examining implicit sociolinguistic perception should be selected and results interpreted in light of the type and time course of the perceptual processes of interest. The elicitation conditions (Preston, 2010) under which sociolinguistic knowledge is accessed can shape the associations that are revealed in a given task, measure, or moment. While all of these measures access some implicit listener representation that links linguistic and social information, the amount of processing time and type of response elicited differ greatly, and these differences may account for the divergence in effects among the measures. Form-meaning links found in a task that requires one type of processing (i.e., fast, automatic eye movements) do not necessarily imply that the same links will arise in a task that requires the other type (i.e., social evaluations), and vice versa. Though these tasks almost certainly involve processes that are intertwined and interactive with one another, the details of which are a fruitful area for further research, we cannot assume that results from any given measure of implicit sociolinguistic knowledge will mirror results in another measure. Maintaining this contrast both theoretically and methodologically expands the examination of sociolinguistic knowledge beyond the metalinguistic (stereotype) versus implicit (marker) boundary and allows us to capture the dynamic nature of online sociolinguistic processing with respect to time and automaticity.

Additionally, the contrast between effects of the valley girl persona and effects of the macrosocial designation of the Californian raises questions about the relevance of different kinds of social information at varied stages of processing. The valley girl persona figures here in the more controlled and introspective measures, while the macrosocial characterization influences the early and more automatic responses. Do these effects indicate truly separate cognitive links between TRAP-backing and these different kinds of social information, accessible via different degrees of control in processing? Or does this contrast reflect different expressions of the same cognitive representation? It may be the case that a link between TRAP-backing and California generally is stored separately from a link between TRAP-backing and the valley girl, and that different measures tap into these separate representations. Another possibility is that the valley girl persona and her associated linguistic style are embedded within a listener's social representation of Californian origin. Since origin is a broader social designation that could conceivably encompass any number of personae and their associated linguistic styles, listeners perhaps access the mapping between Californian origin and backed TRAP more rapidly upon hearing the speech signal,

with a longer time frame required to narrow the feature's association with a particular characterological type of Californian. When listeners have the time with which to access this persona, as in a social evaluation task, they can then more confidently tie TRAP-backing to a valley girl style than to broad Californian origin, as the broader designation may encompass linguistic styles that do not prominently include this feature. However, if the link between the valley girl and TRAP-backing is enveloped within a cognitive representation of general Californian, this would predict that both Californian origin and the valley girl persona would be linked with backed TRAP in more controlled evaluations. Here, TRAP-backing did not influence Californian origin ratings, even for listeners from this region. This suggests that the valley girl persona may overlap with, but not be encompassed completely by, the broader Californian designation—a linguistic feature like TRAP-backing can therefore be evaluated as part of a valley girl style without requiring the speaker to be from California.[8]

The means by which representations of macrosocial versus persona-based representations are formed cognitively could perhaps explain the contrast observed here and provides an additional area for future work. While exemplar theoretic approaches have made clear the significance of repeated episodic experience with linguistic features to future processing (e.g., Johnson, 2006; Pierrehumbert, 2001), representations linking social and phonetic information can also be formed through events that can directly create or strongly reinforce an ideological expectation (Drager & Kirtley, 2016), like a parodic performance or a metalinguistic discussion. The valley girl persona is enregistered by name via parodic performances, popular discourse, and other stereotypical products, while Californian speech styles may be encoded via repeated experience with native Californian speakers. I do not wish to assert that macrosocial California-ness is unrelated to the stereotypical means by which ideologies of the valley girl and other personae are enregistered, nor do I wish to claim that encounters with parodic depictions of a character type do not constitute "experience." However, the weighting of sociolinguistic features encountered via ideologically loaded stereotypes and via interactional exposure may perhaps relate to how and when sociolinguistic effects emerge in processing. Future work may explore the many means by which implicit sociolinguistic expectations can be created and how this bears on different stages of sociolinguistic perception.

NOTES

1. A non-Californian speaker was used in this experiment to eliminate the possibility that the speaker would be heard as Californian, or as a valley girl, regardless of the vowel quality produced. A free-response survey was conducted via Amazon's Mechanical Turk, examining listener impressions of the speaker reading sentence-long productions. The speaker was heard as white in all but 1 of 141 responses, her perceived age averaged 32, and her perceived location of origin was split among the U.S. East Coast, Midwest, and West.

2. One duplicate worker identification entry was found, for which both entries were removed. Listeners who spent over 1 year living outside of the United States between the ages of 5 and 18 years were eliminated from analysis. Language background was also collected following the main task. One participant who did not self-report their native language as English was removed from the dataset.

3. The analysis of experiment 1 employs logistic and linear regression models without random effects, while the analysis of experiment 2 employs mixed-effects models. This is because the structure of the

data in each experiment differs—for experiment 1, the random effects of participant and item included in the models in experiment 2 are not relevant, as listeners heard an entire concatenated list and provided one response set. In experiment 1, there were thus only four items (backed and front crossed with fillers and no fillers). Both of these by-item differences are tested as fixed effects, and participants responded only once per measure in experiment 1. Thus, neither can be included as random intercepts.

4. Model fit was tested stepping up from a model including only fixed effects of fillers and backness to test the influence of each possible participant factor and interactions between significant factors, through chi-square comparisons of the sums of the squares of the residuals (Baayen, 2008).

5. Results from a subset of conditions in experiment 2 were presented in D'Onofrio (2015b), analyzing word choice and gaze fixations from the entire duration of processing, not including the early stage analyzed here.

6. Data from an additional five of the original participants who did not self-report as natively American English speaking, and/or whose average reaction times were longer than 3 sec, and/or whose eye gaze was not accurately recorded in all parts of the screen, were removed prior to analysis.

7. While this type of analysis does not allow for a subtler assessment of eye gaze's evolution over the course of each trial, here the effect of interest was the coarse effect of each social prime on reactions in the earliest time window, for comparison with other measures in this paper. The nuanced ways in which social information modulates gaze over the entire course are an important area for future research.

8. While the present study cannot indicate whether these effects illustrate overlapping or altogether separate cognitive representations, future work might test whether completely orthogonal or even contradictory sociolinguistic associations arise at different stages of processing.

REFERENCES

Allopenna, Paul D., Magnuson, James S., & Tanenhaus, Michael K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38:419–439.

Baayen, R. Harald. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Babel, Anna M. (2016). *Awareness and control in sociolinguistic research*. Cambridge: Cambridge University Press.

Bargh, John, & Chartrand, Tanya. (1999). The unbearable automaticity of being. *American Psychologist* 54:462–479.

Becker, Kara, Aden, Anna, Best, Katelyn, & Jacobson, Haley. (2016). In B. Evans, V. Fridland, T. Kendall, & A. Wassink (eds.), *Publication of the American Dialect Society: Speech in the Western states*. Vol. 101. Durham: Duke University Press. 107–134.

Boersma, Paul, & Weenink, David. 2011. Praat: Doing phonetics by computer, version 5.2.17. 1992–2011. Available at: http://www.praat.org. Accessed May 24, 2016.

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

——. (2012). The implicit association test and sociolinguistic meaning. *Lingua* 122:753–763.

Dahan, Delphine, Drucker, Sarah J., & Scarborough, Rebecca A. (2008). Talker adaptations in speech perception: Adjusting the signal or the representations? *Cognition* 108:710–718.

Devine, Patricia G. (1989) Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56:5–18.

Donald, Kevin, Kikusawa, Ritsuko, Gaul, Karen, & Holton, Gary. (2004). Language. In J. Goggans & A. Difranco (eds.) *The PacificrRegion: The Greenwood encyclopedia of American regional cultures*. Westport: Greenwood Press. 281.

D'Onofrio, Annette. (2015a). Perceiving personae: Effects of social information on perceptions of TRAP-backing. *University of Pennsylvania Working Papers in Linguistics* 21:31–39.

——. (2015b). Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics* 19:241–256.

D'Onofrio, Annette, Eckert, Penelope, Podesva, Robert J., Pratt, Teresa, & Van Hofwegen, Janneke. (2016). The low vowels in California's Central Valley. In B. Evans, V. Fridland, T. Kendall, & A. Wassink (eds.), *Publication of the American Dialect Society: Speech in the Western states*. Vol. 101. Durham: Duke University Press. 11–32.

Drager, Katie, & Kirtley, M. Joelle. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In A. Babel (ed.), *Awareness and control in sociolinguistic research*. Cambridge: Cambridge University Press. 1–24.

Driscoll, Anna, & Lape, Emma. (2015). Reversal of the Northern Cities Shift in Syracuse, New York. *University of Pennsylvania Working Papers in Linguistics* 21:41–47.

Evans, Jonathan St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review* 13:378–395.

Giles, Howard. 1970. Evaluative reactions to accents. *Educational Review* 22:211–227.

Hay, Jennifer, Warren, Paul, & Drager, Katie. (2005). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34:458–484.

Hinton, Leanne, Moonwomon, Birch, Bremner, Sue, Luthin, Herb, Van Clay, Mary, Lerner, Jean, & Corcoran, Hazel. (1987). It's not just the valley girls: A study of California English. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 13:117–128.

Huettig, Falk, Rommers, Joost, & Meyer, Antje S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica* 137:151–171.

Johnson, Keith. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34:485–499.

Johnstone, Barbara, Andrus, Jennifer, & Danielson, Andrew E. (2006). Mobility, indexicality and the enregisterment of "Pittsburghese." *Journal of English Linguistics* 34:77–104.

Kahneman, Daniel, & Frederick, Shane. (2002). A model of heuristic judgment. In K. Holyoak & R. Morrison (eds.), *The Cambridge handbook of thinking and reasoning*. Cambridge: Cambridge University Press. 267–294.

Kahneman, Daniel, & Tversky, Amos. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3:430–454.

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

Koops, Christian, Gentry, Elizabeth, & Pantos, Andrew. (2008). The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics* 36:91–101.

Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William, Ash, Sharon, & Boberg, Charles. 2006. *The atlas of North American English: Phonetics, phonology, and sound change*. New York: Mouton de Gruyter.

Lambert, Wallace, Hodgson, Richard, Gardner, Robert, & Fillenbaum, Samuel. (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology* 60:44–51.

Levon, Erez, & Fox, Sue. (2014). Salience and the sociolinguistic monitor: A case study of ING and TH-fronting in Britain. *Journal of English Linguistics* 42:185–217.

Lieberman, Matthew D. (2007). Social cognitive neuroscience: a review of core processes. *Annual Review of Psychology* 58:259–289.

Payne, B. Keith, Burkley, Melissa A., & Stokes, Mark B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology* 94:16–31.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee and P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 137–157.

Plichta, Bartlomiej. (2013). Akustyk: Speech analysis and synthesis plug-in for Praat. Available at: http://github.com/akustyk. Accessed October 1, 2012.

Podesva, Robert. J., Hall-Lew, Lauren, Brenier, Jason, Starr, Rebecca, & Lewis, Stacy. (2012). Condoleezza Rice and the sociophonetic construction of identity. In J. M. Hernandez-Campoy & J. A. Cutillas-Espinosa (eds.), *Style-shifting in public: New perspectives on stylistic variation*. Amsterdam: John Benjamins. 65–80.

Pratt, Teresa, & D'Onofrio, Annette. (2017). Jaw setting and the California Vowel Shift in parodic performance. *Language in Society* 46:283–312.

Preston, Dennis. (1996). "Whaddayaknow?": The modes of folk linguistic awareness. *Language Awareness* 5:40–74.

———. (2010). Variation in language regard. In E. Zeigler, P. Gilles, & J. Scharloth (eds.), *Variatio delectate: Empirische Evidenzen und theoretische Passungen sprachlicher Variation*. Frankfurt am Main: Peter Lang. 7–27.

Rayner, Keith. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124:372–422.

Schneider, Walter, & Shiffrin, Richard M. (1977). Controlled and automatic human information processing I: Detection, search and attention. *Psychological Review* 84:1–66.

Schnoebelen, Tyler, & Kuperman, Victor. (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija* 43:441–464.

Squires, Lauren. (2013). It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17:200–237.

———. (2016). Processing grammatical differences: Perceiving versus noticing. In A. M. Babel (ed.), *Awareness and control in sociolinguistic research*. Cambridge: Cambridge University Press. 80–103.

Staum Casasanto, Laura. (2008). *Experimental investigations of sociolinguistic knowledge*. Ph.D. dissertation, Stanford University.

Sumner, Meghan, Kim, Seung-Kyung, King, Ed, & McGowan, Kevin. 2014. The socially-weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Language Sciences* 4:article 1015.

Tanenhaus, Michael K. (2007). Spoken language comprehension: Insights from eye movements. In G. Gaskell (ed.), *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press. 309–326.

Villarreal, Daniel. (2016). *The construction of social meaning: A matched-guise investigation of the California Vowel Shift*. Ph.D. dissertation, University of California, Davis.