Personae and phonetic detail in sociolinguistic signs

A N N E T T E   D ' O N O F R I O

Northwestern University, USA

A B S T R A C T

Social meaning-based approaches to linguistic variation treat variation as a semiotic system, in which sociolinguistic signs—indexical links between linguistic forms and social meanings—serve as interactional resources that individuals use to project personae. This article explores the perceptual nature of the links between social personae and linguistic forms, examining how information about a speaker's persona can influence a listener's linguistic perceptions of a continuous phonetic feature. Using a phoneme categorization task, this study examines associations between gradient phonetic manifestations on a continuum from /æ/ to /ɑ/ and three social personae. Findings illustrate that the social persona made relevant for a listener influences the ways in which points on this phonetic continuum are categorized phonemically as either TRAP or LOT. Overall, this shows that the social constructs of personae influence phonetically detailed perceptions of linguistic material. (Sociolinguistic perception, personae, indexicality, sociophonetics, sociolinguistic signs)*

I N T R O D U C T I O N

Sociolinguistic signs

Variationist and ethnographic studies of social meaning in context (e.g. Eckert 2000, 2008a, 2016; Zhang 2005; Podesva 2007; Moore 2010; Bucholtz 2011) have begun to unravel the multiple and complex social meanings that can be associated with the use of a word, an intonation pattern, or a sound. This approach, often called the 'third wave' approach to variation (Eckert 2012), emphasizes that speakers do not use linguistic features solely as a reflection of their social group membership. Individuals constantly exploit linguistic variation's meaningfulness to construct and reconstruct their identities and attitudes in interactional practice. This perspective emphasizes that linguistic features become socially meaningful through their inclusion in styles—combinations of features that are used in social practice (Coupland 2007; Eckert 2008a).

Highlighting the social meaning of linguistic variation necessarily treats sociolinguistic variation as a semiotic system (Eckert 2016), in which linguistic features are imbued with social meanings for both speakers and listeners. Styles may therefore be conceptualized as clusters of sociolinguistic signs, with linguistic forms

coming to signify particular social meanings through the context of their use. These signs—feature-meaning links—are not predetermined one-to-one mappings, nor are they static. The relationship between form and meaning is always dependent on any number of contextual factors, and signs are constantly in flux as they are interpreted and reinterpreted in practice (Silverstein 2003; Eckert 2008a). The ability of linguistic features to flexibly and mutably become associated with a field of shifting meanings is what makes them amenable to use as a social resource, able to accommodate shifting social dynamics. Thus, as resources that are inherently changeable, a crucial property of linguistic forms is that they are underspecified for meaning (Eckert 2016:3). That is, the meaning of a form is only specified when used in a particular interactional moment (Silverstein 1976; Bourdieu 1977; Bakhtin 1981). As Bourdieu notes, 'understanding is not a matter of recognizing an invariable meaning, but of grasping the singularity of a form which exists only in a particular context' (Bourdieu 1977:647). In relation to phonetic features, social and linguistic context necessarily condition the interpretation of acoustic material in a given moment.

Theorizing sociolinguistic signs as interactional resources entails that speakers and listeners themselves maintain cognitive representations of these signs, linking linguistic forms with social meanings (Campbell-Kibler 2011; Drager & Kirtley 2016). Treating signs as cognitive realities requires an examination of what, exactly, each component of a sign looks like in an individual's mind: what kinds of social meanings are linked to what kinds of forms? Further, given the underspecified nature of linguistic features, listeners must have some means of picking out or interpreting the social meaning indexed by a feature in a given instance of its use. This article explores the nature of listeners' links between an interactionally significant social construct—the *persona*—and phonetic forms, asking how personae might perceptually influence linguistic interpretations of continuous acoustic material. Results illustrate that expectations for the phonetic manifestations of phonemic categories can be linked with personae in perception, suggesting that listeners must represent these personae in the mind and tie them to phonetically detailed linguistic categories.

Personae

To consider a linguistic style as inseparable from its social meanings requires an analysis of style as 'persona management' (Coupland 2002:197). Styles emerge as speakers project these personae, and the recognition of styles or registers relies on the recognition of their associations with 'typifiable social personae or practices' (Agha 2005:12). These personae in turn both constitute and reflect the macro-social patterns observable in larger scale studies of linguistic variation. Coupland (2002:198) asserts the mutually constitutive nature of the two: 'It is in relation to group norms that stylistic variation becomes meaningful; it is through individual stylistic choices that group norms are produced and reproduced'.

Studies of style have illustrated the interactional immediacy of personae, relative to macro-social categories (Podesva 2007, 2011; Moore 2012). For example, in Podesva's (2007) study of phonation type in styles used by a gay professional, the speaker uses falsetto to perform a particular persona, the diva, via attributes and stances associated with the persona, like expressiveness. Podesva (2007) illustrates that interactionally, the speaker is not 'performing gayness', but is instead conveying a holistic social type, which may in turn be tied ideologically to gay male identity more broadly construed.

Studies of style as persona-based drive forward the notion that a linguistic feature is imbued with meaning only within its context. Speakers use these features at a particular time, for a particular purpose—they may project one persona in one context, while projecting an entirely different persona in another. This differs from conceptions of linguistic variation as connected statically to macro-social categories, with the use of a linguistic feature falling out as a consequence of a speaker's position in a broader social landscape. Instead, the persona that a speaker projects is fluid and emergent in interaction. Crucially, interactional context involves both the speaker and the listener, as meaning emerges in the intersubjective space between a speaker, who produces the utterance, and the listener, who interprets it (Silverstein 1976; Bourdieu 1977; Bakhtin 1981). Addressing the dialogic nature of language, Bakhtin notes that 'every concrete act of understanding is active: it assimilates the word to be understood into its own conceptual system… and is indissolubly merged with the response' (Bakhtin 1981:282). The study of listener perceptions provides a window into this 'conceptual system', which contributes to and is formed by the interpretation of links between linguistic features and social meanings. The perceptual perspective is particularly important because of the active interpretive work that listeners contribute to the meaningfulness of linguistic variation. Listener interpretations are 'the negotiated evaluation of a speaker's projected persona relative to the local contextualization of the talk, but also relative to listeners' personal experiences and normative expectations' (Coupland 2002:202).

While personae are thus theorized as socially meaningful constructs for both speakers and listeners, studies of interactional personae have tended to focus on the speaker perspective, with analytic techniques extrapolating from context the persona that the speaker aims to project in a given moment. However, for the styles used by speakers to be taken up as socially meaningful (indeed, to become indexical of the projected persona), they must be recognized or interpreted by a listener. Listeners must thus maintain representations of these social personae in the mind, and they must link them with linguistic material cognitively. Yet relatively little work has examined the perceptual nature of personae as connected to linguistic variation (c.f. MacFarlane & Stuart-Smith 2012; Pharao, Maegaard, Møller, & Kristiansen 2014; D'Onofrio 2015). The present article aims to add to our understanding of how social personae are linked cognitively with linguistic material for a listener, asking how a speaker's perceived social persona can influence the phonetic forms that a listener expects from that speaker. In particular, I explore

how a single phonetic continuum (from /æ/ to /ɑ/) can be differently analyzed phonemically, depending on a listener's expectations of the speaker's social persona.

Social meaning and phonetic detail

Researchers in the sociophonetic realm have shown that socially stratified forms can be extremely phonetically detailed, both acoustically (e.g. Di Paolo & Faber 1990; Foulkes & Docherty 1999; Pharao et al. 2014; Drager 2015) and articulatorily (e.g. Lawson, Scobbie, & Stuart-Smith 2013). Furthermore, this work has made possible the study of the continuous nature of phonetic features (e.g. Milroy & Gordon 2003; Gorman & Johnson 2013). Given that linguistic variation can be socially stratified along any number of fine phonetic correlates, the breadth of phonetic nuance needs to be incorporated into models of sociolinguistic production and perception.

Notably, these finding are consistent with the large body of evidence supporting exemplar theoretic approaches to speech perception, which have shown that listeners encode phonetic detail of experienced utterances, as well as speaker information corresponding to these utterances, in the mind (Goldinger 1996; Johnson 1997, 2006; Pierrehumbert 2001, 2016; Foulkes & Docherty 2006; Docherty & Foulkes 2014; Drager & Kirtley 2016). Work in this area has illustrated that listeners do not filter out socioindexical variation in perceiving speech; on the contrary, speaker-specific cues are included as part and parcel of how utterances are encoded cognitively. Thus, our representations of linguistic categories (lexical items or phonemes, for example) are necessarily tied to and conditioned by aspects of the speaker. This raises questions about the nature of the social information that is tied to particular phonetic manifestations of phonemic categories in the mind.

A number of studies have illustrated that fine phonetic cues, perceived in context, can evoke particularized social evaluations (e.g. Campbell-Kibler 2007, 2011; Levon 2014; Pharao et al. 2014), including of the social persona of the speaker. Another strand of work has conversely examined the ways in which social information about a speaker can itself influence linguistic processing, revealing that not only do listeners link phonetic detail with social evaluations, they also must maintain links between this social information and linguistic categories in the mind. This work examining the influence of social information on linguistic processing, however, has nearly always focused on the influence of large scale macro-social groups or speaker categories such as gender (Johnson, Strand, & D'Imperio 1999; Strand 1999), age (Hay, Warren, & Drager 2006; Koops, Gentry, & Pantos 2008; Drager 2011), sexual orientation (Mack & Munson 2012), race (McGowan 2015) or social class (Hay et al. 2006), categories that are typically put forth in this work as nonagentive on the part of the speaker, and relatively static (with the exception of age). As outlined above, however, the interactional use of linguistic forms does not only, and perhaps not most immediately, mark demographic category membership. Speakers project more detailed social types in interactions, updatable in the moment, the nuance of which is not captured in an investigation that

focuses only on larger-scale social categories. Podesva (2006, 2011) has shown that individual speakers use detailed phonetic realizations of features in a socially meaningful fashion (that is, as signs) to project social personae. As an example, Podesva demonstrates that a more acoustically extreme phonetic realization of a variant can intensify a social meaning associated with a less extreme realization, both with regard to intonational contours (2011) and in the case of duration and intensity of stop release bursts (2006). While the link between phonetic detail and personae has been demonstrated to some extent in speaker productions, and other work has shown the link between this detail and macro-social categories in listener perceptions, much less work has tied these strands together to examine whether listeners use social personae when linguistically interpreting phonetic material in perception.

In this article, I approach links between phonetically detailed linguistic categories and social personae from the listener perspective. Specifically, I explore whether (i) information about a speaker's social persona can influence listeners' expectations of boundaries between phonemes, and (ii) whether two unrelated social personae can be perceptually associated with the same phonemic boundary shift, indicating the multiplicity of social personae that can lead to similar phonetic expectations. I focus in particular on perceptions of a backed TRAP vowel, which has been separately associated with two distinct personae: the Valley Girl (Hinton, Moonwomon, Bremner, Luthin, Van Clay, Lerner, & Corcoran 1987; D'Onofrio 2015; Pratt & D'Onofrio 2017) and the Business Professional (Podesva, Hall-Lew, Brenier, Starr, & Lewis 2012; D'Onofrio 2019). I ultimately show that primed expectations of the Valley Girl or the Business Professional persona both lead listeners to classify backer tokens as TRAP phonemes than they would without these primes, or than they would with a persona-based prime associated with a phonemic boundary shift in the opposite direction (the Chicago Bears Fan, a persona characterized in part by a following of the National Football League's Chicago team, the Bears). This indicates that listeners maintain and perceptually deploy cognitive links between phonetic manifestations of the phoneme TRAP and the Valley Girl and Business Professional personae. Notably, however, the shifts in perceptual categorization for each of these two personae arise at different points on an acoustic continuum of backness. The range of phonetic forms within the phoneme TRAP with which these personae are associated are thus overlapping, but distinct from one another. This suggests that listeners not only maintain socially specific phonemic categories as determined by macro-social expectations like age (Drager 2011) or gender (Strand 1999); they also form them based on expectations of more particular social personae. I therefore argue that models of the type of social information tied to instantiations of phonemic categories in the mind should be expanded to include personae, and that such models would benefit generally from a more nuanced view of what constitutes social information than is afforded by macro-social categories alone.

METHODS

TRAP-backing

The sociolinguistic patterning of /æ/, or the TRAP vowel, often referred to as *short a*, has featured prominently in studies of US regional dialects. Work on Californian speakers has found that nonprenasal TRAP is backing and lowering in apparent time, in contrast with patterns of TRAP raising and fronting in other areas of the US, as in the Northern Cities Shift (e.g. Eckert 2000; Labov, Ash, & Boberg 2006). TRAP-backing has been documented in work on the California Vowel Shift (CVS) since at least the 1980s (Hinton et al. 1987). Within the state, the CVS, including the backing of TRAP, has been found to be most advanced in the urban coastal centers of San Francisco (Eckert 2008b) and Los Angeles (Kennedy & Grama 2012), though occurring to a robust degree throughout the less-urban, inland areas of the state as well (Podesva, D'Onofrio, Van Hofwegen, & Kim 2015; D'Onofrio, Eckert, Podesva, Pratt, & Van Hofwegen 2016). Work in the Western dialect region (Labov et al. 2006) more broadly has illustrated that TRAP-backing is prevalent not only in California, but also in other Western states like Oregon (Becker, Aden, Best, & Jacobson 2016) and Washington (Wassink 2015) as well.

In relation to this regional patterning, TRAP-backing has been associated with a Californian persona: the Valley Girl. This social type is characterized as young, female, wealthy, shallow, and materialistic. Though typically viewed as associable with the Southern Californian San Fernando Valley for which she is named, the label has been expanded to describe young, female social types outside of this region as well, nearly always with negative connotations (Donald, Kikusawa, Gaul, & Holton 2004; Bucholtz, Bermudez, Edwards, Fung, & Vargas 2007). Work analyzing parodic performances of this character type, such as Whoopi Goldberg's imitation of a Valley Girl in her 1984 stand-up comedy show 'The spook show' (Hinton et al. 1987) and the more recent *Saturday night live* skit series *The Californians* (Pratt & D'Onofrio 2017), as well as work examining perceptual associations (D'Onofrio 2015) has found that both in comedic enactments of this persona and in listener perceptions, a backed TRAP vowel is linked ideologically with the Valley Girl persona.

TRAP-backing has also been associated with a persona quite distinct from the Valley Girl and not clearly associated with California at all—the Business Professional. Studies of context-based intraspeaker variation (Podesva et al. 2012) and of listener perceptions (D'Onofrio 2019) have shown that the backed TRAP variant indexes formality, intelligence, and professionalism in what has been operationalized as a Business Professional persona. While the origins of this indexical association are less clear-cut than for the Valley Girl, since they are not as overtly performed in exaggerated and named mass-media depictions of the social type, two possible indexical orders may be responsible. First, this association may

have arisen via American ideologies of British English. Received Pronunciation (RP) maintains a phonemic split between the TRAP class of words, which was historically pronounced with the fronter /æ/ vowel, and the BATH class, pronounced with low back /ɑ/ (Wells 1982). As American English does not have this distinction, all words in the BATH class are pronounced with the same vowel as TRAP in American varieties, such that the distinction could be heard as a generalized reanalysis of TRAP as a LOT-like vowel /ɑ/. BATH seems to be a particularly salient feature to American ears, being associated with British English. American imitations of British English tend to include this feature. This association between a backed TRAP vowel (or an RP BATH vowel heard as a TRAP vowel) and British English is situated within broader American ideologies of British English (e.g. Boberg 1999). Studies of Americans' reactions to British English found that Americans rated British English speech samples as more intelligent, confident, successful, and ambitious than American English samples (Stewart, Bouchard, & Giles 1985). A general sense of high status and greater education that may index aristocracy in the British context may have transmuted via indexical orders (Silverstein 2003) to index highly successful white-collar professionals in the American context.

Second, links between TRAP-backing and intelligence, professionalism, and formality may have stemmed from ideological oppositions to regional TRAP-*raising* in a number of locations of the United States. Together, the presence of a relatively raised nonprenasal TRAP vowel in speakers of the Inland Northern Cities (e.g. Eckert 2000; Labov et al. 2006), the South (Thomas 2005), and in some nonprenasal environments in the *short a* splits of the Northeastern cities of New York and Philadelphia (Labov et al. 2006), may have attached social meanings, particularly for speakers of regional dialects in which TRAP does not raise. Many of these regional dialects are stigmatized in the wider American popular imagination, with the stereotypical New Yorker, Philadelphian, Northern Cities speaker, or Southerner tending to be regarded as speakers of some form of marked or nonstandard English (Preston 1999). Bolstering the potential for TRAP-raising to be classified as a stigmatized variable is recent dialectological work showing an apparent time movement away from the feature in these regions, particularly in more educated or supralocally oriented young speakers (Prichard & Tamminga 2012; Driscoll & Lape 2015; Wagner, Mason, Nesbitt, Pevan, & Savage 2015). It may therefore be the patterns of negative evaluation of, and movement away from, localized TRAP-raising that are responsible for the Business Professional meaning of TRAP-backing. If oppositional to the Northern Cities Shift, this indexical meaning should entail a link not only with the backing of TRAP in the vowel space, but also with its lowering and monophthongization.

Notably, I take the respective links between the Valley Girl and Business Professional personae and TRAP-backing as rooted in ideological notions of these personae and their linguistic styles, rather than in a listener's exposure to 'actual' Valley Girls or Business Professionals (however we might characterize 'actual' instantiations of these personae). It is unclear, for example, that any individual speaker would

self-identify as a 'Valley Girl', and while more may self-identify as 'Business Professionals', individuals do not embody this professional persona at all times, by necessity. Instead the personae investigated here serve as ideological abstractions associated with a particular set of recruitable social meanings that can be indexed through the use of a linguistic style, of which TRAP-backing is a part. As personae themselves are constructs drawn upon by speakers toward particular interactional ends (Eckert 2016), they are not meant to represent social groups or categories of which a speaker can themselves be a member—a divergence from the ways in which speakers have typically been conceived of as embodying a racialized or gendered category or an age group. Instead, the meaningful conception of these personae comes from ideological notions of what characterizes these contrasting types, and these characterizations are recruited to project social meanings through the use of linguistic features that index those types. A listener's ideas about the phonetic features associable with a persona thus come not only from meta-pragmatic sources like enregistered performances of, or meta-discourse about, that persona (e.g. Pratt & D'Onofrio 2017), but also from the interactional moments in which these personae are invoked stylistically (e.g. Podesva et al. 2012).

That two unrelated personae (the Valley Girl and the Business Professional) can be indexed by a similar feature in linguistic production—backed TRAP—is inherent to theories of indexicality that pose the sociolinguistic sign as multiplex and mutable (Silverstein 2003; Eckert 2008a). Using a phoneme categorization task, I examine whether these social personae can shift perceptual classifications of phonetic tokens along a continuum of backness in ways that reveal their shared association with TRAP-backing. Further, I examine how different acoustic ranges of TRAP-backing may be linked with these respective personae in listener perceptions.

Phoneme categorization

This study employs a phoneme categorization task common in investigations of sociolinguistic perception. In this paradigm, researchers provide listeners with top-down social information about a speaker. They then examine the effect that this information has on classification of a token as one phoneme or another. Social information about the speaker, or the social prime, can be presented explicitly, in writing (Niedzielski 1999; Strand 1999) or photographs (Johnson et al. 1999; Hay et al. 2006), or implicitly, via cues in the speech signal itself (Drager 2011) or placement of a cue in the experimental environment (Hay & Drager 2010). After being presented with social information, listeners are asked to categorize a series of tokens on a continuum between two phonemes (/s/ and /ʃ/, for example). Analysis examines whether the social information modulates the perceived boundary between the two phonemes. For example, Strand (1999) found that listeners who thought they were hearing a female voice placed the boundary between /s/ and /ʃ/ at a higher acoustic frequency (closer to the /s/ end of the continuum) than listeners who thought they were hearing a male voice. In other words,

an ambiguous token on such a continuum is more likely to be heard as /ʃ/ if the speaker is assumed to be a woman, but /s/ if the speaker is assumed to be a man, reflecting a listener's sociolinguistic expectations related to speaker gender—on the aggregate, women are more likely to produce fricatives in a higher overall frequency range than men.

The vowel categorization task presented here investigates whether listeners show a link between social personae and phonetic manifestations of TRAP when categorizing points on a continuum from /æ/ to /ɑ/. Following previous work using such tasks (Strand 1999; Hay et al. 2006), this experiment provides listeners with explicit top-down social information about a speaker, then examines how this information affects categorization of a given vowel token. On a continuum from from /æ/ to /ɑ/, ambiguous tokens between the two poles can be analyzed as either a relatively backed production of TRAP, or as a relatively fronted production of LOT. Here, I employ three persona-based social primes to characterize the speaker. First, I include Valley Girl and Business Professional primes. Given these personae's associations with TRAP-backing in production, I expect that listeners who receive these social primes will be more likely to analyze a given continuum token as TRAP (rather than LOT), even when that token is relatively backed.

To serve as a point of contrast, I also include a social prime expected to yield the opposite effect, leading listeners to be more likely to analyze ambiguous tokens as LOT, rather than TRAP. I thus select a social prime associable with *LOT-fronting*: the 'Chicago Bears Fan'. As a hallmark feature of the Northern Cities Shift, LOT-fronting has been found in speakers from the Northern Cities of the United States, including Chicago (Labov et al. 2006; McCarthy 2011). Stereotypical depictions of fans of Chicago's NFL football team, the Bears, have been pervasive in both local (Hallett & Hallett 2014) and more widespread US media—a well-known series of *Saturday night live* skits parody this persona in the characters of 'Bill Swerski and the Superfans' who famously deploy the Northern Cities Shift when discussing 'da Bears'. For this experiment, I use the Chicago Bears Fan persona as a counterpoint to the Valley Girl and Business Professional. It is important to note, however, that the stereotypical depiction of a Chicago Bears Fan is an older male. While this was deployed as an easily nameable persona associated with LOT-fronting, its ideological association with the Northern Cities Shift may be tempered by the fact that the voice is a young female's. The inclusion of this persona is intended only to serve as a point of comparison to the Valley Girl and Business Professional personae, and the meanings and instantiations of this persona in particular are not a major focus of this study. I hypothesize that overall tokens in the Bears Fan condition will be more frequently categorized as LOT than those in the Valley Girl and Business Professional conditions.

Finally, I include a condition that provides listeners with no prior social information, to serve as a baseline against which the other conditions are compared. In studies of sociolinguistic perception, analysis typically involves contrasts between different social conditions for which divergent effects are expected

(here, the contrast between the Bears Fan on one hand, and the Valley Girl and Business Professional on the other). However, when comparing oppositional social conditions, it may be the case that the observed contrast is not driven by both social primes equally (Campbell-Kibler 2011). The use of a baseline condition permits comparisons of social primes against the absence of an induced social prime, rather than against one another, which allows the influence of each social prime to be considered on its own. Notably, the baseline condition is not to be interpreted as in any way devoid of social information—every utterance crucially contains a wealth of social indices, and cues from a speech signal interact with top-down expectations that listeners bring to each utterance regardless of explicit a priori information. Inclusion of a baseline condition simply provides responses that are derived from interpretations of the signal independent from an explicitly induced social expectation. Thus, all three social primes will initially be compared with the Baseline condition in analysis, with statistical models treating the Baseline condition as the factor level default.

Stimuli

Stimuli were created through resynthesis of naturally produced tokens, which were elicited from word-list readings from one speaker. The author, a female native speaker of American English in her mid-twenties at the time of recording, born and raised in the Midwestern United States, recorded productions of a series of TRAP-LOT minimal pairs read from a word list. Of these productions, three of the minimal pairs were selected based on clarity of the tokens: *black-block*, *map-mop*, and *sack-sock*. These tokens were selected because they showed clear, monophthongal bands in the spectrogram without interruption from glottalization or other disturbances in the signal. For each pair, the naturally produced TRAP and LOT words were used as poles to create a resynthesized continuum between /æ/ and /ɑ/. These tokens were created using the Akustyk package (Plichta 2013) in Praat (Boersma & Weenink 2011). The command 'Create Continuum' in Akustyk takes two vocalic tokens and digitally manipulates an original token's dynamic formant structure at a number of incremental time steps throughout the vowel (in this case, at ten millisecond intervals) to create a continuum from one token to the other. Tokens on a continuum of vowel quality were created by manipulating the original TRAP token's first three formant values in nine equal steps to progressively match the original LOT token's. For the *black-block* continuum, the /l/ preceding the vowel was included in the resynthesis, such that the continuum produced resynthesized tokens of /l/+vowel, which were then spliced into the original *black* /b/-/k/ frame. For *sack* and *black*, vowels were segmented from the onset of voicing as observable in the spectrogram and waveform (at the nearest zero crossing). For *map*, the vowel was segmented at the boundary between the nasal and the vowel, visually determined by a sharp increase in amplitude, and taken at the closest zero crossing following this boundary. For all three words, the ending boundary of

TABLE 1. *Midpoint F1 and F2 values for continuum steps used in the phoneme categorization task.*

Continuum step	BLACK F1	BLACK F2	MAP F1	MAP F2	SACK F1	SACK F2
1	993	1904	1002	2098	1042	1815
2	995	1828	1006	1979	1034	1772
3	995	1757	1016	1896	1038	1710
4	995	1681	1044	1784	1029	1612
5	990	1605	1057	1735	1014	1579
6	986	1554	1079	1647	1016	1559
7	983	1492	1089	1521	1021	1487
8	975	1419	1097	1398	1008	1405
9	964	1364	1121	1289	997	1383

the vowel was determined as the offset of voicing, as measured by the end of the F2 band in the spectrogram, including the closure.

The manipulation was then performed via Akustyk on the extracted vowels, producing nine steps between the TRAP and LOT tokens. The re-synthesized tokens match the mean duration of the two original productions, with pitch and formant values above F3 matching the original TRAP token. The script resynthesizes vowel tokens to 10 KhZ, 16 bits. After resynthesis, tokens were spliced into the original preceding and following consonantal frames from the TRAP token, with the exception of the preceding fricative in SACK, for which a corresponding fricative continuum from the /s/ preceding the original SACK token to the /s/ preceding the original SOCK token was created via Praat script and spliced to match its corresponding vowel token. All stimuli were then scaled for peak amplitude in Praat. Formant measurements for the resulting stimuli are provided in Table 1.

Tokens spanned from at least 1800 Hz to at least 1400 Hz in F2. In a sample of speakers from California's Central Valley, where the CVS has been shown to be in progress, women between ages eighteen and thirty-five show a mean TRAP F2 value of 1680 Hz, as compared to women over sixty-five who show a mean TRAP F2 value of 1800 Hz (measurements calculated on dataset from D'Onofrio et al. 2016). For these same sets of speakers, the younger women show a mean LOT F2 value of 1280 Hz, while the older women show a mean LOT F2 of 1300 Hz. The continua used in the present experiment thus represent a range from productions of /æ/ (similar to older, nonshifted Californian speakers) to productions of /ɑ/ (slightly fronter than the /ɑ/ tokens produced by female Californian speakers). Both F1 and F2 ranges varied among the continua, as shown in Table 1. To ensure that tokens were heard as equally ecologically valid, all were subjected to naturalness ratings on a sliding scale from 0 ('sounds not at all manipulated') to 100 ('sounds digitally manipulated or synthesized') via Amazon's Mechanical Turk. Each of the twenty-seven total tokens (three word pairs x nine steps) was rated by ten raters (270 total raters). Overall, all three word pairs received similar naturalness ratings averaged across the nine-token continuum (*black* tokens Mean = 43; *sack* tokens = 37.5;

map tokens = 46). Furthermore, no individual token's mean rating deviated more than half a standard deviation from the overall word pair's mean rating, and no clear patterns in ratings were observable based on location on the continuum. This indicates that while listeners did hear the tokens as somewhat manipulated (though slightly closer to the 'not at all manipulated' pole), none of the tokens were heard as significantly more manipulated or synthesized than the others.

Design and procedure

Each participant was given one of four possible social primes.[1] In the Baseline condition (i), no social information was provided about the speaker. Listeners in the Valley Girl condition (ii) were told, 'The speaker you will hear has been described as a Valley girl'. Listeners in the Business Professional condition (iii) were told, 'The speaker you will hear has been described as a business professional'. Listeners in the Chicago Bears condition (iv) were told, 'The speaker you will hear has been described as a Chicago Bears fan'. Listeners in the persona prime conditions (ii)–(iv) were provided with this corresponding sentence in writing two times prior to the main task. The information was first presented in an introductory screen, accompanied by task instructions. Then, listeners completed four practice trials to become accustomed to the task. Practice trials contained end-point TRAP and LOT tokens that were not members of any of the three target minimal pairs, produced in the same voice as in the main trials (*jack-jock, stack-stock, lack-lock, tap-top*). The practice trials thus gave listeners exposure to the speaker prior to the main task, aiming to acquaint the listener with the speaker's endpoints of TRAP and LOT. Poles of the minimal pairs in the practice task were chosen to provide listeners with experience using the keyboard to make their choices, but responding to stimuli that were unambiguous. Following the practice round, and prior to the main task, listeners were once again provided with written condition-based social information. The same voice was used in both the practice and main tasks because it allowed for two exposures to the written social information about the speaker, before the practice round and before the main task. While exposure to the voice in the practice round may have provided context for responses in the main task, since every participant completed the same practice trials in the same order, any cross-condition or cross-word effects of this early exposure would be controlled.

Both social prime and word pair were manipulated between subjects. Each participant heard every step on ONE of the three word-pair continua (*black-block, map-mop,* or *sack-sock*) one time, for a total of nine critical trials per participant. A given listener thus provided categorizations of each token for the entire continuum for one minimal pair, with analysis controlling for word pair and measuring the effect of a priori social information on how steps were categorized phonemically. In each trial of the practice and main tasks, listeners were shown a screen with the word choices (e.g. *sack* or *sock*) on the left and right side of the screen, respectively, with left-right placement of the word randomized by participant. In each trial of the main task,

listeners heard one of the steps on a given word-pair continuum. Presentation order of trials was randomized for each participant. The auditory token played at the beginning of each trial, and participants then used a [1] (left) or [0] (right) on their keyboards to indicate which word they heard (e.g. *sack* or *sock*). The experiment advanced to the next trial when the listener responded, or after five seconds if no response was provided, eliciting a relatively speeded response to the stimulus.

Participants

All participants were recruited and compensated via Amazon's Mechanical Turk (AMT), a web interface that crowd-sources individuals to complete tasks online for payment. The use of crowd-sourced participants for experimental research via online interfaces has become increasingly common, not only in linguistic research generally (Snow, O'Connor, Jurafsky, & Ng 2008; Callison-Burch & Drezde 2010; Schnoebelen & Kuperman 2010; Sprouse 2011), but also in the area of speech processing specifically (see Eskénazi, Levow, Meng, Parent, & Suendermann 2013 for an overview). In sociolinguistic perception studies, online participants have frequently been used to obtain social evaluations of phonetically detailed stimuli in matched-guise style tasks (e.g. Campbell-Kibler 2007; Levon 2014). The major benefits of using crowd-sourced online platforms are the ability not only to recruit and compensate large numbers of registered participants quickly, but also to yield a more diverse demographic sample than is typically obtained in university laboratory settings (Buhrmester, Kwang, & Gosling 2011). Comparisons of validity across lab and online settings in fact show that patterns obtained in a lab setting can be replicated online as reliably as, and sometimes more reliably than, in a lab setting (Schnoebelen & Kuperman 2010; Burhmester et al. 2011; Crump, McDonnell, & Gureckis 2013;). In the realm of phonetics and phonology specifically, a growing number of studies have successfully deployed phonetically detailed speech perception tasks through AMT including discrimination, recognition, classification, and identification in noise of phonetically detailed stimuli (e.g. Yu & Lee 2014; Walker & Campbell-Kibler 2015; Kimball & Cole 2016; Denby, Schecter, Arn, Dimov, & Goldrick 2018). Yu & Lee (2014) in particular compare in-lab participant performance to AMT worker performance on a two-alternative forced-choice word-categorization task for phonetic continua—the same task used in the present study—as well as AXB style discrimination tasks, to assess influences on categorization of sibilant continua. They find that the two data sources show consistent results with one another, offering support for the valid use of online speech-perception studies. Notably, few of these studies have used online experiments to test the influence of social primes on speech perception, a methodological area for future research and further validation.

Certainly, drawbacks also accompany using crowd-sourced data, namely in the degree of control and monitoring allowed the researcher. To address these issues in the present study, a number of checks were put into place in the experimental

design, recruitment, and analysis to verify that the experiment was being completed as requested. First, participants were filtered by reported location in their user profile—only those who reported that they were residents of the United States, and were verified via IP address, were able to perform the task. Second, participants were provided with a sound check before they were allowed to proceed, to ensure that they would in fact be listening to the stimuli in the task. In the sound check, listeners were required to listen to an auditory word and number, and to type the word and number they heard in order to proceed to the task. This test consisted of a voice not used in the experiment. Finally, participants were required to complete the entire experiment before they were given a unique authentication code, which they were required to enter into an interface in Mechanical Turk. Data was used only from those workers who accurately entered their assigned code.

360 participants were recruited: thirty participants in each of the twelve word-pair social-prime combinations (three word pairs x four social primes). Given that the experiment was conducted online, using a large number of participants who completed only nine trials each (responding to each step of one word-pair continuum only one time) allows for the inclusion of a wider demographic sample of listeners without inflating the number of response tokens analyzed. Given that fewer controls can be put into place with an online experiment, as compared to one conducted in a lab setting, the use of many participants for fewer trials each also attenuates the possibility that participants reduce their attention throughout a longer task, and it greatly reduces the statistical influence of any one participant on the patterns observed. From an initial screen on Mechanical Turk, participants were led to an external web experiment through which they completed the task.

Participant information was collected in a questionnaire following the main task, which elicited self-reported participant age, gender, language(s) spoken natively, locations lived, and the ages at which they lived there. Listener location of origin was coded for broad dialect region according to the Atlas of North American English (Labov et al. 2006). Further, given the prominence of TRAP-backing in the California Vowel Shift throughout the Western dialect region, as well as the location-based association between the Valley Girl persona and the state of California (within the West), speakers were also placed into a binary category of Western origin versus non-Western origin, as measured by whether or not they lived in a state designated within the boundaries of the West (Labov et al. 2006) from birth to age eighteen. Participants who did not report English as a native language, did not grow up in and live in the United States at the time of the experiment, or reported having a hearing disorder were eliminated from analysis, as were all responses from any listeners who completed the task more than once. Type of listening device (speakers, headphones, or earbuds) was tested as a potential predictor of response, but was not found to influence results. While listener productions have been shown to correspond to phonemic perceptions (Hay et al. 2006), production data for participants was not collected due to the online nature of the task, but the correspondence between the effects observed and listener productions of these vowels

TABLE 2. *Participant self-reported background information for participants included in final dataset, by social prime condition.*

Social prime	Mean age	% Western US origin	% female	% male	% another gender	Total N
Baseline	31	27%	47%	53%	0%	75
Chicago Bears	36	16%	41%	56%	2.7%	73
Business Professional	28	23%	38%	62%	0%	81
Valley Girl	31	21%	39%	61%	0%	83

would be an interesting avenue for future study. Again, the inclusion of a large number of different participants and the collection of social correlates that may predict their exposure to and productions of certain vowels aim to attenuate the effect of any individual listener's ability to drive patterns in the data. Since the procedure moved on to each trial after five seconds with or without a response, those who were missing more than two of the nine target categorization responses were eliminated from analysis, as they likely experienced technical difficulties or were not fully attending to the task. A total of 312 participants remained in the final dataset used for analysis. Some of the remaining participants did not provide responses to one or two trials (an overall total of twenty-four missing responses), which yielded a total token count of 2,784 in the final dataset.

Of the participants whose data was usable, the mean participant age was 32.41% of participants self-reported as female, 58.4% as male, and 0.6% as another gender. 22% of participants grew up in the Western United States, with the remainder split between Northeastern US origin (23%), the South (24%), the North (16%, encompassing the Inland North and North Central dialect regions) and the Midlands (10%). 6% lived in multiple dialect regions prior to age eighteen, and they were placed in a separate 'multiple regions' category. Participant background broken down by social prime is summarized in Table 2.

As this table shows, participant background was fairly evenly balanced across the four social primes, and, as noted in the results section below, the only participant factor that showed a significant influence on categorization results was binary Western versus non-Western origin, which is fairly evenly distributed across social primes, with the Chicago Bears prime showing fewer Western participants than the others.

RESULTS

Mean LOT responses (as opposed to TRAP responses) for each continuum step are plotted in Figure 1 by social prime. Word pair did not significantly interact with social prime condition or step (nor was the three-way interaction significant), as

described in further detail below, indicating that social prime effects emerged controlling for the word pair assessed. Thus, Figure 1 collapses responses across word pairs for clarity. Each pane shows a comparison between responses from participants with the Baseline (no information) prime, in gray, as compared to those who received a persona-based prime, in black.

Given the associations between manifestations of TRAP and the personae discussed above, it was predicted that the Business Professional and Valley Girl primes would occasion fewer LOT responses (more TRAP responses) than the Baseline and Bears primes, patterns borne out in the results (Figure 1). The Business Professional condition diverged from the Baseline condition overall, illustrating that listeners who were told the speaker was a Business Professional were significantly more likely to respond to a token as TRAP than those with no speaker information. The Valley Girl prime also trended in the expected direction: those who thought the speaker was a Valley Girl were more likely to classify a given token as TRAP as compared to those in the Baseline condition (Figure 1, bottom).

Notably, the Chicago Bears prime did not diverge from the Baseline, with nearly identical response patterns between the two (Figure 1, top). This means that, as expected, listeners with the Business Professional prime and the Valley Girl prime also responded TRAP more frequently than those with the Chicago Bears prime. The very similar categorization results for the Baseline and Chicago Bears primes may indicate that the Chicago Bears Fan persona is not as strongly linked with phonetic manifestations of these vowels as compared with the other personae, at least in this voice, and thus does not lead to strong expectations of phonetic manifestations of phonemic categories. Another possibility is that the tokens contained socioindexical cues aside from vowel quality that aligned more closely with the Chicago Bears Fan persona than with the other persona-based primes, a plausible interpretation given that the speaker was from the Midwestern US. Regardless, the differences between the Chicago Bears social prime and the Business Professional and Valley Girl primes operate in the expected direction—the boundary between TRAP and LOT is perceived to be backer when listeners thought the speaker was a Business Professional or Valley Girl, as compared to a Chicago Bears Fan, or when they had no information about the speaker.

Statistical analysis of the trends observable in Figure 1 was performed using a mixed-effects logistic regression model in R (R Core Team 2013) via the *lmer* function in the *lme4* package (Bates, Maechler, Bolker, & Walker 2014). The models estimated categorical selection of a LOT word (versus TRAP). The fixed effects tested were that of social prime (four levels: Baseline (default), Chicago Bears Fan, Valley Girl, or Business Professional), and word pair (three levels: *black-block* (default), *map-mop*, *sack-sock*) as categorical predictors, and continuum step, as well as continuum step of the immediately preceding trial, as continuous predictors. Interactions among the experimental manipulations (social prime, word pair, and continuum step) were also tested. Participant background effects of gender, age, Western vs. non-Western dialect region, and listening device
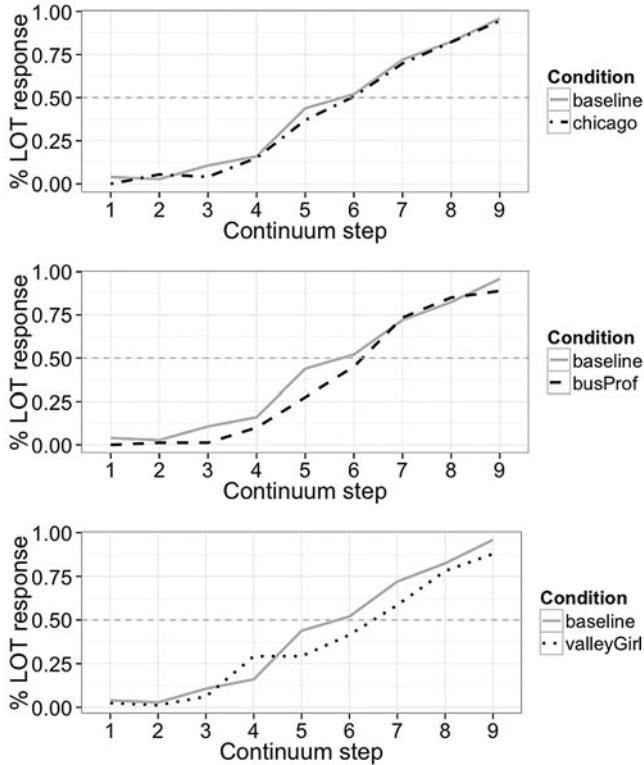
FIGURE 1. Percent LOT categorization by continuum step (1 = front, 9 = back), by condition. Baseline condition in gray, versus Chicago Bears Fan prime (top); Business Professional prime (middle); and Valley Girl prime (bottom).

were also tested as fixed effects. A random intercept of participant was included. Effects were omitted from the final model when they did not serve as statistically significant predictors and did not improve model fit as assessed through a comparison of the sums of the squares of the residuals using the *anova* function in R.

The model summary treats the Baseline condition as the default factor level for the social prime effect, so coefficients reflect the degree of influence relative to the Baseline. The fixed effect of word pair serves as a control variable such that effects of social prime reflect differences over and above the influence of any one word pair, while also making visible what the nature of differences between the word pairs were. Following Drager (2011), the inclusion of the main effect of continuum step treats place on the TRAP-LOT continuum as a single continuous control variable. Again, this allows for any significant differences among social primes found in the model to reflect differences in categorization generalized across the entire continuum. Interactions among these three factors were also tested. An interaction between

continuum step and social prime examines how social primes differently influenced categorization responses at different places on the continuum. In other words, it tested whether the influence of a given social prime on categorization was dependent on how backed the token was. Similarly, the interaction between continuum step and word pair, and between social prime and word pair, assess whether categorization differences among these factors are dependent on one another. A three-way interaction amongst these predictors was also tested. While the presentation order of the continuum steps was randomized across participants, continuum step encountered in the immediately preceding trial was also tested in the model, in case presentation order may have influenced results and was required as a control predictor. As mentioned above, participant social background factors were also tested.

Social prime, continuum step, word pair, and participant regional origin served as significant simple predictors of categorization. A significant interaction between social prime and continuum step also emerged. The other possible interactions between continuum step, word pair, and social prime (including their three-way interaction) did not prove significant nor improve model fit. Preceding step did not show a significant effect on categorization and did not improve model fit, and therefore was not retained in the final model. Similarly, participant factors of gender, age, and listening device did not significantly influence categorization and were not retained. Only the binary participant factor of Western versus non-Western dialect region significantly predicted responses. The summary of the best-fit model is shown in Table 3.

The significant interaction between social prime and continuum step indicates that on a granular level, categorization of steps at particular parts of the continuum were differently influenced by the persona-based primes. In other words, the influence of social prime on categorization of a token as TRAP versus LOT is dependent on location on the continuum, or how backed the token is. While both the Business Professional and Valley Girl primes increased the overall likelihood that listeners will respond to a token on the continuum as TRAP, these effects differ based on location in the continuum. More specifically, the slope of the categorization curve for the Business Professional prime (Figure 1, middle, dashed line) differs significantly from the Baseline (solid gray line), and the slope of the curve for the Valley Girl prime (bottom, dotted line) differs marginally from the Baseline. In addition, a re-leveled model treating Business Professional as the default social prime showed that the interaction between social prime and continuum step differs significantly between the Valley Girl and the Business Professional primes (Valley Girl prime x continuum step interaction: Beta = -0.50, SE = 0.13, Z = -3.81 P = 0.0001), indicating that the influence of each of these primes arose in different areas of the continuum.

The control factor of word pair was a significant predictor of responses, indicating that the place on the continuum at which the boundary between TRAP and LOT fell depended strongly on which word pair was being heard, due to lexical,

TABLE 3. *Model summary for vowel categorization (Baseline as default; N = 2784).*

Predictor	Estimate	Std. error	Z value	P value
(Intercept)	−10.077	0.688	−14.656	<0.0001***
Continuum step	1.374	0.100	13.683	<0.0001***
Condition = *Chicago Bears*	0.019	0.766	0.254	0.799
Condition = *Business Professional*	−2.460	0.852	−2.888	0.004**
Condition = *Valley Girl*	0.788	0.708	1.113	0.266
Word pair = *MAP-MOP*	1.875	0.278	6.744	<0.0001***
Word pair = *SACK-SOCK*	4.221	0.302	14.000	<0.0001***
Listener Western origin = *Western*	−0.702	0.250	−2.815	0.005**
Step x condition = *Chicago Bears*	−0.031	0.124	−0.254	0.800
Step x condition = *Business Professional*	0.281	0.135	2.079	0.038*
Step x condition = *Valley Girl*	−0.214	0.114	−1.885	0.059

phonological, or phonetic differences among the pairs. However, this was only significant as a simple effect—these word-pair differences did not significantly interact with the other fixed effects (including the three-way interaction between social prime, continuum step, and word pair). This indicates that effects of the social prime and continuum step were not specific to any given word pair. The effects can thus be interpreted over and above this word-pair-based difference, for which the model controls.

Comparison of the shapes of these curves illustrates the nature of this significant difference between the Business Professional and Valley Girl primes. Indeed, these two persona-based social conditions did not affect the continuum uniformly across the entire range of backness (Figure 1). Rather than simply shifting the entire categorization curve downward (which would indicate fewer TRAP responses across the board, or no continuum step by social prime interaction), effects of the Business Professional and Valley Girl primes diverged from the Baseline only in some parts of the continuum, while not in others. The Business Professional prime led to a greater likelihood of a TRAP response only in the fronter portion of the continuum, then joining the Baseline and Chicago Bears Fan primes in the backer portion of the continuum. Listeners thus expected a Business Professional to use backed TRAP when backed to a smaller degree; this influence diminishes past a certain point of backness. This contrasts with the Valley Girl prime. While the Valley Girl curve is close to Baseline for the fronter portion of the continuum, it diverges and occasions a greater likelihood of a TRAP response in the backer half of the continuum. This suggests that listeners were more likely to expect very backed TRAP vowels from a Valley Girl than from the other social primes. While both the Business Professional and Valley Girl primes led to greater expectations of TRAP-backing generally in comparison to the Chicago Bears Fan or Baseline primes, these two different persona-based social primes also differed significantly from one another depending on how the token was manifested with respect to phonetic backness.

Table 3 also illustrates that listeners who were raised in the West (lived there their entire lives prior to turning eighteen years old) were more likely to hear a given continuum token as a TRAP word than listeners who were raised in non-Western dialect regions. That is, Westerners' boundary between TRAP and LOT was drawn at a position further back in the vowel space than non-Westerners' boundary. Notably, this effect did not interact with social persona condition. Regardless of the social information provided, Western listeners were generally more likely to expect a backer TRAP than non-Western listeners were.

DISCUSSION AND CONCLUSIONS

This study takes perceptions of TRAP-backing as a case study to explore the ways in which social personae can influence the linguistic processing of speech sounds. Results demonstrate that the same phonetic continuum, from /æ/ and /ɑ/, is divided into phonemic categories differently depending on the persona-based information provided about the speaker. The Valley Girl and Business Professional primes led to a backer perceptual boundary between TRAP and LOT than the Chicago Bears Fan prime did. These listeners thus showed perceptual links between both the Business Professional and the Valley Girl and forms of TRAP-backing, confirming associations observed in production patterns (e.g. Podesva et al. 2012; Pratt & D'Onofrio 2017). This suggests that listeners deploy expectations of how certain personae will produce phonemes when interpreting acoustic information linguistically. Not only do speakers project these holistic personae in interaction then (Zhang 2005; Podesva 2007; Eckert 2008a), listeners also draw upon these constructs in interpreting the speech of others.

Results corroborate a growing body of work in sociolinguistic perception to demonstrate the cognitive reality of sociolinguistic signs: if listeners' linguistic behavior can be shaped by their persona-based expectations, they must maintain links between these social personae and linguistic forms in the mind. Given that the phonemic categorization of an utterance is dependent upon the persona from whom that token emerges, conceptualizing particularized personae as tied to specific manifestations of acoustic material for a listener may more fully reflect how these links exist as constructs used in interactional practice. This suggests a theoretical expansion, beyond macro-social categories, of what can constitute the socioindexical information linked with linguistic representations in the mind. A more nuanced view of the social information tied with phonetic productions not only reflects the way that variation is used semiotically in interaction (Podesva 2007, 2011), it more clearly corresponds with episodic theories of speech perception. Much work supporting exemplar models has illustrated that we derive cognitive categories like words or phonemes from clusters of acoustically detailed episodes, and that we encode the rich phonetic detail offered by episodes of speech (e.g. Goldinger 1996; Foulkes & Docherty 2006; Pierrehumbert 2016). This study suggests that alongside this phonetic detail, rich and interactionally relevant social constructs beyond macro-

social categories—even when rooted in stereotypes—are also represented and linked with linguistic expectations.

Crucially, this expands and elaborates the nature of the social information tied to listener expectations of how phonemic categories are realized phonetically. For example, these findings suggest that social meanings of phonetic forms need not be defined in direct opposition to one another, with a linguistic contrast corresponding to a pre-existing social contrast, as can be implied by the invocation of macro-social categories with a limited set of 'levels' (age-based generations, socioeconomic class levels, binary gender categories, etc.). Instead, different personae altogether, themselves inhabiting these broader macro-social categories, can be linked with distinct, but perhaps overlapping, ranges of vocalic productions. When interpreting speech, then, we can draw upon impressions of the speaker that are more specific and holistic than demographic group membership.

The nature of the task also allows a closer look at how the influence of persona-based speaker information arose in phonetically detailed ways, given that effects of the social primes differed depending on whether participants were responding to more subtly backed tokens or to extremely backed tokens. Underspecification is a crucial property of the sociolinguistic variable, which allows for language to flexibly index a variety of changing social meanings (Eckert 2016). The study presented here reflects this property: two different social personae (the Business Professional and Valley Girl) both shifted listeners' overall responses in a way that indicated an expectation of a backer boundary between TRAP and LOT. However, the significant interaction found between continuum step and social prime suggests that this meaning potential does not apply to all phonetic manifestations of 'backed TRAP' equally. Instead, the influence of a given persona-based prime is a function of how phonetically backed the token in question is. The specificity of the effects observed in this study in fact appears to reflect the ways that different phonetic forms of TRAP-backing are ideologically associated with these respective personae. The Valley Girl persona has been enacted in parodic performances, which commonly involve the hyperbolic use of linguistic features (Coupland 2001). Exaggerated forms of the California Vowel Shift, including extreme TRAP-backing, have been found in performances of Southern Californian social types (Hinton et al. 1987; Pratt & D'Onofrio 2017). Contrastingly, the Business Professional meanings appear to be connected to less extreme realizations of TRAP-backing, likely those more akin to the variance found in intraspeaker context-based style shifts (e.g. Podesva et al. 2012). This speaks to the significance of detailed phonetic form in specifying the persona that a feature can index (Podesva 2007, 2011). The persona-based social meaning primed in this study specifies the linguistic interpretation of ambiguous speech sounds in a phonetically detailed manner consonant with that persona's linguistic style.

Additionally, though not the focus of the present study, the effect of listener Western origin illustrates that a listener's background contributes to perceptions of phonetic forms, following work that has established the listener-dependent

nature of sociolinguistic perception (e.g. Campbell-Kibler 2007; Levon 2014) and work that has shown the link between listener perceptions and productions (e.g. Hay et al. 2006; Sumner & Samuel 2009; Kendall & Fridland 2012). Listeners who were raised in the Western dialect region were more likely to perceive an ambiguous token as backed TRAP. Although Western listeners are certainly exposed to speakers who are likely to produce backer TRAP vowels, listener place of origin not only determines exposure, but also ideologies, and these ideologies can directly affect sociolinguistic perceptions of speech (Levon 2014; Pharao et al. 2014). Future work on regionally implicated features like TRAP-backing might aim to tease apart the way that a macro-social measure like listener region of origin may encompass both everyday exposure to particular phonetic manifestations of a linguistic feature, as well as ideologies or stereotypes derived from meta-pragmatic activity (Silverstein 2003), both of which can influence sociolinguistic perception.

While this study provides a step toward understanding sociolinguistic signs as cognitive constructs, much further work is required to explore how linguistic forms are specified for social meaning for a listener. For example, it should be noted that the results pertaining to TRAP-backing in particular must be interpreted in light of the fact that this voice, in this style, showed these social associations. The findings presented here are contextualized by a number of factors, including the careful citation style in which the stimuli were produced, expectations induced by the experimental context, and other features of the voice itself. The Chicago Bears Fan prime did not differ from Baseline in this task, for example, but we might expect a significant difference in categorization if presented alongside a different voice. The question of how phonetic detail interacts with additional contextual characteristics of the utterance aside from the speaker's persona remains a fruitful area for further study (Pharao et al. 2014). Furthermore, the *persona* is merely one social construct crucial to the conveyance of stylistic meaning. Future work is required to expand our understanding of the ways in which affect, stance, or interactional role can condition linguistic processing. Such work would provide further bridges between our knowledge of how sociolinguistic styles are deployed to convey social meaning in interaction, and how listeners perceive and represent such styles in the mind.

NOTES

[1]Note that a fifth condition was also tested using the same paradigm and stimuli—a 'California' condition testing the influence of macro-social California origin. As the present study takes as its focus the contrast between separate personae, as opposed to a contrast between different types of social information (macro-social versus persona-based), results of the California condition are not included here. However, they do not interfere with the findings and claims made in this article, and all details and analysis of the California condition results can be found in D'Onofrio (2015), as can additional discussion of the ramifications of persona- versus macro-social information on listener perceptions (particularly the contrast between Californian origin and the Valley Girl persona).

REFERENCES

Agha, Asif (2005). Registers of language. In Alessandro Duranti (ed.), *A companion to linguistic anthropology*, 23–45. Malden, MA: Blackwell.

Bakhtin, Mikhail (1981). Discourse in the novel. *The dialogic imagination: Four essays*. Austin: University of Texas Press.

Bates, Douglas; Martin Maechler; Ben Bolker; & Steve Walker (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. Online: http://CRAN.R-project.org/package=lme4.

Becker, Kara; Anna Aden; Katelyn Best; & Haley Jacobson (2016). Variation in West Coast English: The case of Oregon. In Valerie Fridland, Tyler Kendall, Betsy E. Evans, & Alicia Beckford Wassink (eds.), *Speech in the western states. Vol. 1: The coastal states. Publication of the American Dialect Society*, 107–34. Durham, NC: Duke University Press.

Boberg, Charles (1999). The attitudinal component of variation in American English foreign (a) nativization. *Journal of Language and Social Psychology* 18(1):49–61.

Boersma, Paul, & David Weenink (2011). Praat: Doing phonetics by computer. Version 5.2.17. 1992–2011. Online: http://www.praat.org.

Bourdieu, Pierre (1977). The economics of linguistic exchanges. *Social Science Information* 16:645–68.

Bucholtz, Mary (2011). *White kids: Language, race and styles of youth identity*. Cambridge: Cambridge University Press.

———; Nancy Bermudez; Lisa Edwards; Victor Fung; & Rosalva Vargas (2007). Hella Nor Cal or totally So Cal? The perceptual dialectology of California. *Journal of English Linguistics* 35:325–52.

Buhrmester, Michael; Tracy Kwang; & Samuel D. Gosling (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspective on Psychological Science* 16(1):3–5.

Callison-Burch, Chris, & Mark Dredze (2010). Creating speech and language data with Amazon's Mechanical Turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, 1–12.

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

——— (2011). The sociolinguistic variant as a carrier of social meaning. *Language Variation and Chang* 22:423–41.

Coupland, Nikolas (2001). Dialect stylization in radio talk. *Language in Society* 30:345–75.

——— (2002). Language, situation and the relational self: Theorizing dialect-style in sociolinguistics. In Penelope Eckert & John R. Rickford (eds.), *Style and sociolinguistic variation*, 185–210. Cambridge: Cambridge University Press.

——— (2007). *Style: Language variation and identity*. Cambridge: Cambridge University Press.

Crump, Matthew J. C.; John V. McDonnell; & Todd M. Gureckis (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE* 8(3):e57410.

Denby, Thomas; Jeffrey Schecter; Sean Arn; Svetlin Dimov; & Matthew Goldrick (2018). Contextual variability and exemplar strength in phonotactic learning. *Journal of Experimental Psychology: Learning, Memory and Cognition* 44:280–94.

Di Paolo, Marianna, & Alice Faber (1990). Phonation differences and the phonetic content of the tense-lax contrast in Utah English. *Language Variation and Change* 2:155–204.

Docherty, Gerard J., & Paul Foulkes (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142:42–56.

Donald, Kevin; Ritsuko Kikusawa; Karen Gaul; & Gary Holton (2004). Language. In Jan Goggans & Aaron Difranco (eds.), *The Pacific region: The Greenwood encyclopedia of American regional cultures,* 281. Westport, CT: Greenwood Press.

D'Onofrio, Annette (2015). Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics* 19:241–56.

——— (2019). Sociolinguistic signs as cognitive representations. In Lauren Hall-Lew, Emma Moore, & Robert J. Podesva (eds.), *Social meaning and linguistic variation*. Cambridge: Cambridge University Press, to appear.

———; Penelope Eckert; Robert J. Podesva; Teresa Pratt; & Janneke Van Hofwegen (2016). The low vowels in California's Central Valley. In Valerie Fridland, Tyler Kendall, Betsy E. Evans, & Alicia Beckford Wassink (eds.), *Speech in the western states. Vol. 1: The coastal states. Publication of the American Dialect Society*, 11–32. Durham, NC: Duke University Press.

Drager, Katie (2011). Speaker age and vowel perception. *Language and Speech* 54:99–121.

——— (2015). *Linguistic variation, identity construction and cognition*. Berlin: Language Science Press.

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

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

Eckert, Penelope (2000). *Linguistic variation as social practice*. Oxford: Blackwell.

——— (2008a). Variation and the indexical field. *Journal of Sociolinguistics* 12:453–76.

——— (2008b). Where do ethnolects stop? *International Journal of Bilingualism* 12:25–42.

——— (2012). Three waves of variation study: The emergence of meaning in the study of variation. *Annual Review of Anthropology* 41:87–100.

——— (2016). Variation, meaning, and social change. In Nikolas Coupland (ed.), Sociolinguistics: Theoretical debates, 68–85. Cambridge: Cambridge University Press.

Eskénazi, Maxine; Gina-Anne Levow; Helen Meng; Gabriel Parent; & David Suendermann (eds.) (2013). *Crowdsourcing for speech processing: Applications to data collection, transcription and assessment*. Chichester: John Wiley & Sons.

Foulkes, Paul, & Gerard J. Docherty (1999). Sociophonetic variation in 'glottals' in Newcastle English. In John J. Ohala, Yoko Hasegawa, Manhari Ohala, Daniel Granville, & Ashley C. Baily (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, 1037–1040. Berkeley: University of California, Berkeley.

———, & ——— (2006). The social life of phonetics and phonology. *Journal of Phonetics* 34:409–38.

Goldinger, Stephen (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory & Cognition* 22:1166–83.

Gorman, Kyle, & Daniel Ezra Johnson (2013). Quantitative analysis. In Robert Bayley, Richard Cameron, & Ceil Lucas (eds.), *Oxford handbook of sociolinguistics*, 214–40. Oxford: Oxford University Press.

Hallett, Jill, & Richard Hallett (2014). 'Just a regular guy': Dialect variation and parodic stylization on Chicago radio. Paper presented at New Ways of Analyzing Variation 43, Chicago, Illinois.

Hay, Jennifer, & Katie Drager (2010). Stuffed toys and speech perception. *Linguistics* 48:865–92.

———; Paul Warren; & Katie Drager (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34:458–84.

Hinton, Leanne; Birch Moonwomon; Sue Bremner; Herb Luthin; Mary Van Clay; Jean Lerner; & Hazel Corcoran (1987). It's not just the valley girls: A study of California English. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 117–28.

Johnson, Keith (1997). Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.), *Talker variability in speech processing*, 145–65. San Diego, CA: Academic.

——— (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34:485–99.

———; Elizabeth A. Strand; & Mariapaola D'Imperio (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27:359–84.

Kendall, Tyler, & Valerie Fridland (2012). Variation in perception and production of mid front vowels in the U.S. *Southern Vowel Shift*. Journal of Phonetics 40:289–306.

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

Kimball, Amelia, & Jennifer Cole (2016). Pitch contour shape matters in memory. *Proceedings of the International Conference on Speech Prosody* 8:1171–75.

Koops, Christian; Elizabeth Gentry; & Andrew Pantos (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(14):91–101.

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

Lawson, Eleanor; James M. Scobbie; & Jane Stuart-Smith (2013). Bunched /r/ promotes vowel merger to schwar: An ultrasound study of Scottish sociophonetic variation. *Journal of Phonetics* 20:101–110.

Levon, Erez (2014). Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society* 43:539–66.

MacFarlane, Andrew E., & Jane Stuart-Smith (2012). 'One of them sounds sort of Glasgow Uni-ish': Social judgements and fine phonetic variation in Glasgow. *Lingua* 122:764–78.

Mack, Sara, & Benjamin Munson (2012). The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype. *Journal of Phonetics* 40:198–212.

McCarthy, Corinne (2011). The Northern Cities Shift in Chicago. *Journal of English Linguistics* 39 (2):166–87.

McGowan, Kevin B. (2015). Social expectation improves speech perception in noise. *Language and Speech* 58(4):502–21.

Milroy, Lesley, & Matthew Gordon (2003). *Sociolinguistics: Method and interpretation*. Oxford: Blackwell.

Moore, Emma (2010). The interaction between social category and social practice: Explaining was/were variation. *Language Variation and Change* 22:347–71.

——— (2012). The social life of style. *Language and Literature* 21:66–83.

Niedzielski, Nancy A. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Social Psychology* 18:62–85.

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* 43:1–31.

Pierrehumbert, Janet B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*, 137–57. Amsterdam: John Benjamins.

——— (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics* 2:33–52.

Plichta, Bartlomiej (2013). Akustyk: Speech analysis and synthesis plug-in for Praat. Online: http://github.com/akustyk.

Podesva, Robert J. (2006). *Phonetic detail in sociolinguistic variation*. Stanford, CA: Stanford University dissertation.

——— (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11:478–504.

——— (2011). Salience and the social meaning of declarative contours: Three case studies of gay professionals. *Journal of English Linguistics* 39(3):233–64.

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

———; Annette D'Onofrio; Janneke Van Hofwegen; & Seung Kyung Kim (2015). Country ideology and the California Vowel Shift. *Language Variation and Change* 27:157–86.

Pratt, Teresa, & Annette D'Onofrio (2017). Jaw setting and the California Vowel Shift in parodic performance. *Language in Society* 46(3):283–312.

Preston, Dennis (1999). *Handbook of perceptual dialectology 1*. Philadelphia, PA: John Benjamins.

Prichard, Hilary, & Meredith Tamminga (2012). The impact of higher education on Philadelphia vowels. *University of Pennsylvania Working Papers in Linguistics* 18:87–95.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Online: http://www.R-project.org/.

Schnoebelen, Tyler, & Victor Kuperman (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija* 43(4):441–64.

Silverstein, Michael (1976). Shifters, linguistic categories, and cultural description. In Keith Basso & H. A. Selby (eds.), *Meaning in anthropology*, 11–56. Albuquerque: University of New Mexico Press.

——— (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23:193–229.

Snow, Rion; Brendan O'Connor; Daniel Jurafsky; & Andrew Y. Ng (2008). Cheap and fast – but is it good? Evaluating non-expert annotations for natural language tasks. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, 254–63.

Sprouse, Jon (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavioral Research Methods* 43(1):155–67.

Stewart, Mark A.; Ellen Bouchard Ryan; & Howard Giles (1985). Accent and social class effects on status and solidarity evaluations. *Personality and Social Psychology Bulletin* 11:98–105.

Strand, Elizabeth (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18:86–99.

Sumner, Meghan, & Arthur G. Samuel (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60:487–501.

Thomas, Erik R. (2005). Rural white southern accents. In Edgar W. Schneider (ed.), *Varieties of English: The Americas/Caribbean*, 87–114. Berlin: Mouton de Gruyter.

Wagner, Suzanne Evans; Alexander Mason; Monica Nesbitt; Erin Pevan; & Matt Savage (2015). Reversal and re-organization of the Northern Cities Shift in Michigan. Paper presented at New Ways of Analyzing Variation 44, Toronto, Canada.

Walker, Abby; & Kathryn Campbell-Kibler (2015). Repeat what after whom? Exploring variable selectivity in a cross-dialectal shadowing task. *Frontiers in Psychology* 6:546.

Wassink, Alicia Beckford (2015). Sociolinguistic patterns in Seattle English. *Language Variation and Change* 27:31–58.

Wells, John (1982). *Accents of English 3: Beyond the British Isles*. Cambridge: Cambridge University Press.

Yu, Alan C. L., & Hyunjung Lee (2014). The stability of perceptual compensation for coarticulation within and across individuals: A cross-validation study. *The Journal of the Acoustical Society of America* 136:382.

Zhang, Qing (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society* 34:431–66.

Address for correspondence:
Annette D'Onofrio
Linguistics Department
Northwestern University
2016 Sheridan Road
Evanston, IL 60208, USA
donofrio@northwestern.edu