Language in Society **50**, 29–51.
doi:10.1017/S0047404520000871

Affect and iconicity in phonological variation

ANNETTE D'ONOFRIO AND
PENELOPE ECKERT

Northwestern University, USA
Stanford University, USA

ABSTRACT

The study of iconic properties of language has been marginalized in linguistics, with the assumption that iconicity, linked with expressivity, is external to the grammar. Yet iconicity plays an essential role in sociolinguistic variation. At a basic level, repetition and phonetic intensification can intensify the indexicality of variables. Iconicity plays a further role in variation in the form of sound symbolism, linking properties of sounds with attributes or objects. Production studies have shown some phonological variables exhibiting sound symbolism, particularly in the expression of affect. In some cases, the observation of sound symbolism has been largely interpretive. But in others, stylistic variability as a function of speaker affect has provided empirical evidence of iconicity. This article examines the role of iconicity and performativity in transcending the limits of reference, reviews iconicity in production studies, and provides experimental evidence that sound symbolism influences how listeners attribute affect to linguistic variation. (Variation, iconicity, affect)

INTRODUCTION

> We constantly poach the forest of wild meanings in order to improve the strain of domestic ones. (Bolinger 1989:2–3)

The issue of iconicity comes and goes in linguistics, but the field continues to focus on language as a referential system composed of pure symbols, and iconicity continues to appear as a quirky peripheral phenomenon. What appears to be at issue in the marginalization of iconicity in linguistics is its expressivity, a property eschewed in the Enlightenment focus on rationality that dominates not only linguistics but modern ideology more generally. Yet as Jakobson (e.g. 1960) repeatedly pointed out, expressivity is an essential function of language. The study of variation, building onto referential grammars, has also overlooked expressivity, and especially emotivity, focusing on more objective social categories. Yet variation is at the heart of expressivity in language. In this article, we provide production and perceptual evidence of iconicity and affective expression in variation, and consider

© The Author(s) 2020. Published by Cambridge University Press 0047-4045/20 $15.00

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

what iconicity tells us about the nature and role of variation in a theory of language. We begin with a discussion of the limits of language and of reference, and the role of indexicality and performativity in transcending these limits. We then move on to the semiotics of, and the role of iconicity in, variation.

LINGUISTIC BOUNDARIES

The limitation of linguistics to the referential system was built into structuralism, witnessed by de Saussure's (1916) sign as an arbitrary pairing of form and meaning. For de Saussure, the sign is a token in the set of abstract relations that constitute structure, and does not connect directly with referents in the world. While de Saussure recognized iconicity in language, he viewed it as a resemblance in the real world, outside of the sign, and not in the grammar (Joseph 2015). This separation of the grammar from the world continues in linguistics, and the boundary around the grammar is central to much analytic practice, and by extension to disciplinary policing of what is and isn't linguistics. This was explicitly articulated with the coining (Trager 1958) of the term *paralanguage* for meaningful vocal (primarily prosodic) resources that could not be accounted for in a referentially based grammar. Definitions of paralanguage vary, reaching into semiotics more generally, to gesture and even clothing (Pennycook 1985). Ladd (1996) carves out reference for linguistic intonation and assigns noncategorical intonation to paralanguage, which he sees as carrying emotion and attitude. Bolinger (1989:2), by contrast, views all intonation as iconic and attitudinal, arguing that even the intonational difference between question and statement 'is an affective matter (curiosity vs. confidence, etc.), not a "propositional" (grammatical) one'. Experimental studies of intonation support Bolinger's view. Jeong's (2016) study of the perception of polar interrogatives as questions, invitations, or commands shows a conventional association between affective prosodic meanings and illocutionary force. And her study (Jeong 2018) of rising declaratives shows that the steeper the slope of a rising contour, the more likely it is to be heard as inquisitive rather than assertive. So if there is a boundary between the continuous and the categorical, or the affective and the referential, it is fluid, and the mechanism of this fluidity is without question an essential part of language.

Our interest in the affective role of variation grows out of our understanding of variation as a system of indexical signs. Early work in variation viewed the meaning of variation as a reflection of its progress through the socioeconomic hierarchy (Labov 1972). In other words, variables had meaning only by association with their users, what Silverstein (1976) calls *presupposing nonreferential* indexes. This view is adequate if 'the social' is limited to replicable, particularly macrosocial, categories that articulate enduring and widespread hierarchies. But if class, age, gender, and racial hierarchies appear stable, the discourses in which they are maintained and contested are continually changing, and sociolinguistic variation

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

is a central element in this process, as a semiotic resource with which speakers position themselves in an ever-changing social world (Eckert 2019a). The only circumstances under which linguistic differences could be tied directly to macrosocial structure would be if contact, as Bloomfield (1933) claimed, were the only factor in the diffusion of change. But that train has long left the station and explanations for macrosocial variation routinely invoke persona management. A not infrequent pattern of gender divergence in the working class has commonly been attributed to working class men's avoidance of female-led change (e.g. Tagliamonte & D'Arcy 2009). Whether or not this precise claim is true, it recognizes that variation invokes ideology. But it does not follow this insight to its source. If working class men are semiotically distancing themselves from women, it is not because they fear that someone will think they are women, but that someone will think they are LIKE women in some specific and currently salient way. They may be avoiding sounding weak, or gay, or flamboyant—or trying to sound powerful or straight or down-to-earth. Studies that focus on macrosocial distinctions generally recognize that qualities like these are at issue, but treat the categories as set packages and the finer meanings as incidental. The following discussion is based in a view of variation as part of the always ongoing semiosis in which meanings like 'snobbish' or 'weak' are not side effects of macrosocial patterns, but the level at which variation engages the social.

Qualities like snobbishness and weakness are not fixed but historically contingent, gaining meaning in their associations with, for example, class and gender. These associations come about as individuals collectively fashion selves within the constraints of place in the social order. Studies in linguistic anthropology emphasize the ideological construction of forms of personhood—of the subservient and flamboyant Griot and the laconic Noble (Irvine 2001; Gal & Irvine 2019), the cultivated, nonlocal RP-speaker (Agha 2003), the fierce and the femme drag queen (Calder 2019). These forms of personhood embody powerful ideologies at every level. We locate the study of variation within this semiotic tradition, focusing on phonological variation as the auditory detail of the semiotics of personhood. The Griot and the Noble are enregistered figures of personhood, which constitute landmarks in the social-semiotic landscape because they embody a deeply cultural ideological distinction. And the stylistic features that distinguish their speech (e.g. patterns of amplitude, speech rate, pitch range) are fractally available for the expression of components of that distinction within categories (Gal & Irvine 2019). Social change is marked by changes in figures of personhood: flappers, hobos, hippies, valley girls, hip-hoppers all embody social distinctions specific to time and place. If a figure of personhood is a landmark, *persona* is the individual's continual self-presentation with respect to the landscape, and the locus of gradual social change. Persona is built up of detail, combining semiotic resources indexing not only demographic characteristics but stance, affect, qualities, status. These can all be expressed referentially, but style combines them into a construction whose meaning is difficult to articulate, and layers that meaning onto the explicit

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

message content. This is not only economical but strategic, for putting something in an assertion assumes intention and all its attending consequences, while expressing it stylistically opens up a safe space for experimentation and interpretation (Eckert 2019b).

THE SEMIOTICS OF VARIATION

Sociolinguistic variation is entirely performative and indexical. Performativity and indexicality are basic to the grammar more generally, but because they link language to the world, they come into their own in variation. Performativity is the power of a speaker to make change in the world through language. Austin's (1962) canonical performatives involve highly conventionalized acts—promising, threatening, conferring civil and official status. These acts constitute a limited number of culturally salient transactions whose success resides in the speaker, in their personal and situational legitimacy. Austin concluded that all utterances are performative, but one could say that some are more performative than others, based on the extent to which the performativity constructs not only the context but the speaker themselves (Eckert 2019b). Potts' analysis of expressives like *damn!* or *you bastard you!* locates the meaning in the speaker's affect—as he puts it (Potts 2007:180), 'Expressive content is performative in this sense: quite generally, the act of uttering an expressive is the emotive performance'. Butler (1993) takes performativity beyond language, integrating performativity into the social order as people become gendered through the performance of gendered acts. Lodged in a theory of language, Austin's felicity conditions on speech acts point to, but do not engage with, the structures that confer legitimacy (e.g. one must be legally authorized to perform a marriage). Butler's (1993) performativity confers legitimacy as people become gendered through the performance of culturally legitimated gendered acts. And in turn, the accumulation of gendered performances over the population reproduces the very structures that underlie legitimation. We locate variation in this larger performative realm, as stylistic practice creates a persona in the moment—what one might consider the ultimate in performativity.

Sociolinguistic variables are indexical signs, each with a field of potential meanings (Eckert 2008) which become specified as they enter into constructions, or styles. It is not surprising that linguistic anthropologists adopt Peirce's (1868) theory of the sign because unlike de Saussure, Peirce locates signs in the world. Peirce's sign is discursive, consisting of form (the *representamen* or *sign vehicle*), its *object* in the world, and an *interpretant*, or construal of the sign in the moment. The interpretant gives rise to a new sign, making meaning always emergent, a continuous chain of construal, and inseparable from the continuous chain of change in the world. The interpretant construes not only the object, but the nature of the link between the sign vehicle and the object: a *symbolic* link is purely conventional, an *iconic* link is a resemblance, and an *indexical* link is a real-world association.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

An index points out into the world, conducing the listener to search common ground for its object. Linguists mostly focus on indexicals such as determiners, tense markers, or pronouns, which reach into the common ground of the speech situation to fill in the semantics. These are what Silverstein (1976) calls *referential* and *presupposing* indexes. When indexicals point beyond the immediate context, they require—and invoke—personal, communal (Clark 2006) and ideological (Acton 2019) common ground. The demonstrative in an utterance like *there goes that Donald Trump!* indexes stance, and invites the listener to interpret—and perhaps to share—that stance (Acton & Potts 2014). If the listener suspects that the speaker has recently become a Trump fan, they may interpret the use of *that* as indexing affection or amusement rather than distance or contempt. Silverstein (1976) calls these *creative* or *performative* indexicals. Similarly, a form that is phonologically marked in some way (e.g. an unexpectedly creaky vowel), moves the hearer—consciously or unconsciously—to search common ground for a change in the speaker's stance or persona and for local construals of creaky voice. The patterned accumulation of such episodes over time and social-situational space creates a semiotic chain (Agha 2003), constructing a meaning for creak and potentially in the long term a new figure of personhood. It is no doubt through the situated accumulation of stylistic features, among them creaky voice, that the Valley Girl emerged in the seventies as a landmark of White affluent California personhood in the social landscape. In this way, stylistic practice is fundamentally ideological, simultaneously representing and transforming meaning in the landscape.

ICONICITY

An icon is a sign whose sign vehicle is construed as resembling its object in some diagrammatic or sensual way. Iconicity is conventional but it invokes the natural. Iconicity is found at all levels of language, from the diagrammatic iconicity of discourse and syntax (e.g. Haiman 1980) to the more sensual iconicity of sound symbolism. We are concerned here with sound symbolism, which permeates the lexicon, prosody, systems of ideophones (e.g. Bolinger 1950, 1985; Dingemanse 2018) and, we argue, variation. But we begin with one fundamental, diagrammatic device—the intensification of meaning through repetition. Repetition can intensify the denotation (as in *he's a bad bad man*), the implication (as in *run! run! run!*) or the expressivity (as in *damn! damn! damn!*). Expressive intensification can also be achieved through intensification of the sign vehicle, such as phonetic lengthening ([bæ:d mæ:n, rʌ:n, dæ:m]) or fortition ([d'æm]). By building intensification into the sign, repetition and phonetic intensification are at the heart of the performativity of variation, intensifying the social meaning of a sociolinguistic variant. One might say that iconic intensification underlies the functioning of variation itself, which is why the study of variation is a fundamentally quantitative enterprise: the more negative concord a Detroit suburban teenager uses the more burned-out they sound (Eckert 2000). And the phonetic quality that sets apart a sociolinguistic variant

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

can itself be intensified: the more one retracts /æ/ the more one sounds like a Valley Girl (D'Onofrio 2015).

Our primary focus here is on variants whose phonetic form is construed as an acoustic reflection of some aspect of the speaker's persona or subjectivity. Gal & Irvine (2019; Irvine & Gal 2000) offer an account of rhematization, or the iconic construal of an index, as an ideological process. Qualia—sense perceptions such as smoothness, darkness, harshness—are commonly associated with human qualities and from there to features of language (Gal 2013). An icon involves a relation to something in virtue of some kind of resemblance, but the resemblance itself is constructed. It is the ideological moment that brings about rhematization. Phonetic iconicity can reach a metapragmatic level, as in the common attribution of imagined cultural qualities to languages or dialects on the grounds that they are 'nasal' or 'guttural'. It is often said that Beijing Mandarin is 'oily', based on its distinctive rhoticity (Zhang 2005). There is no natural link between rhoticity and oiliness, or between oil and social characteristics, but there is a widespread idea of oiliness as a social quality involving smoothness or unctuousness. Zhang (2005, 2008) documents the association of rhotacization in Mandarin with a well-established Beijing characterological figure—the 'Smooth Operator'—judged metaphorically to be a smooth talker, or 'oily'. Thus, what begins as an index (i.e. rhoticity indexes Beijing origins) undergoes rhematization, or reconstrual as iconic (i.e. rhoticity points to the oiliness of a Smooth Operator). In this way, rhoticity enters a new order of indexicality (Silverstein 2003), from indexing place to indexing a salient association with people from that place. This association did not come out of nowhere. Beijing rhoticity and the image of the Smooth Operator came into semiotic play at a time of social change, as China was opening up economically, and Yuppies emerged as a new breed of players in the international market. The Yuppies, or young professionals working in the new international financial sector, constructed a crisp cosmopolitan persona, including a speech style significantly lacking in rhotacization in contrast with their peers in the state-owned financial sector. Zhang (2018) attributes this stylistic practice not simply to a move away from the old local sphere, but to a larger ideological process, the reopening of social differentiation, 'undoing the commonness' of the cultural revolution.

The power of iconic signs lies in the appearance that they are natural, emanating from the speaker's essence. Iconicity plays a particularly important role in the expressive realm (e.g. Jakobson & Waugh 2002), where the sound attaches performatively to the speaker's body and persona (Eckert 2019b). Ohala (1994), followed by Gussenhoven (2016), have posited biological origins for a range of iconicities. Ohala's (1994) *frequency code*, by far the more robust, builds on the relation between acoustic frequency and the size of the emitting body. Not only are larger vertebrates' voices naturally lower, vertebrates of all sizes use lower pitched vocalizations to appear larger, and higher pitch to indicate appeasement (Morton 1994). These effects find their way into language in a complex web of associations with fundamental frequency and with the acoustic frequency of segments. The latter

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

has been well known since Sapir's (1929) early minimal pair experiments in which participants judged a nonsense syllable with the nucleus [a] to denote a large object, the same syllable with [i] to denote a small one. The acoustic frequency of vowels and consonants finds its way into size-related meanings in the lexicon, in augmentative-diminutive systems (e.g. Silverstein 1994), and in physical (Woodworth 1991) and even social (Silverstein 1994), distance. Both F0 and vowels extend into the more synaesthetic realm, to contrasts of darkness and brightness, heaviness and lightness (e.g. Fónagy 1983; Hubbard 1996), roundness and sharpness (e.g. Köhler 1929; Ramachandran & Hubbard 2001; Maurer, Pathman, & Mondloch 2006; D'Onofrio 2014), and negative and positive affect (Fónagy 1983; Silverstein 1994). These associations apply in music (Gabrielsson & Lindström 2010) and in bodily motion as well (Casasanto & Dijkstra 2010). The heterogeneity of associations with acoustic frequency is a testimony to synaesthesia and the continual interpretive capabilities of human existence.

The frequency of F0 meets the social directly in the size difference between children and adults, and to a lesser extent between adult women and men. By extension, it evokes the power differential that size creates not only between but within age and gender groups, including social power as in the Wolof Griot and Noble styles (Irvine 2001; Gal & Irvine 2019). Certainly the difference in pitch and in physical and social power is greater between children's and adults' voices than between adult male and female voices, but the more commonly cited indexicality is adult gender, where size difference takes on a deeply cultural significance well beyond any biological basis. The bottom line is that iconicity comes into play in the service of ideological hot spots. Biology has been traditionally invoked to justify contested social arrangements such as racial oppression and the subjugation of women, while adults' power over children is not as widely contested. As Gal & Irvine (2019) emphasize, when we see iconicity at work, we have every reason to look for the ideological force that motivates it. At the same time, the frequency code is a vast network of associations, and it would be a mistake to assume that one indexicality—such as gender—precedes the others.

ICONICITY AND EXPRESSIVITY IN VARIATION

Haeri (1996) brought the frequency code specifically into the realm of sociolinguistic variation in her work on Cairene Arabic. Finding that women led in the strength of palatalization and men led in pharyngealization, she noted that across community studies of sound change, women tend to lead in changes involving fronting while men lead in backing. Gordon & Heath (1998) followed up with a more general proposal, linking it explicitly to the frequency code. Haeri emphasized that this pattern interacts with class, specifically in the quality of refinement that is associated with both gender and class and, one might consider, with delicacy and lightness. The interaction of gender with other social factors is illustrated in the case of /s/ in

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

English and at least some other languages. On average, the center of gravity of women's and girls' /s/ has been found to be significantly higher than men's and boys' from an early age (Flipsen, Shriberg, Weismer, Karlssson, & McSweeny 1999) and across cultures (Bekker 2007; Stuart-Smith 2007; Li 2017). This pattern more clearly reflects socialized rather than physiological distinctions, as acoustic frequency of the turbulence of /s/ is only slightly affected by the size of the vocal tract. Additional gender differences are achieved by spreading/rounding the lips and fronting/backing the tongue tip. This gender difference is exploited by drag queens (Calder 2019), and the fractal binary that folds sexual orientation into gender leads to the common association of /s/ fronting with gay men (for a review, see Levon, Maegaard, & Pharao 2017). As in the case of Arabic palatalization and pharyngealization, /s/ can also index class (Bekker 2007; Stuart-Smith 2007), urbanity vs. rurality (Bekker 2007; Podesva & Van Hofwegen 2016; Bekker & Levon 2017), and even relative threat of stance (Holmes & Levon 2017) and fierceness (Calder 2019).

A wide range of affect is not likely to appear in sociolinguistic interviews, and it is difficult for a sociolinguist to catch affective changes in the wild. Yet there have been examples. Eckert (2011) compared the speech of a preadolescent girl in two conversations, one in which she talked mostly about things she liked, another that she initiated in order to complain about what jerks boys are. The nuclei of the LOT and PRICE vowels moved significantly to a higher and backer position in the first conversation than in the second one. Geenberg (2014) asked ten adults to talk to a plush toy pig, first praising her for her cuteness, and then soothing her for a hurt knee. Not only did all ten speakers use a significantly higher f0 in the 'cute piggy' task than in the 'poor piggy' task, they also showed a significantly higher f2 for many of their vowels in the 'cute piggy' task, and none showed a lower f2 in that task for any of the vowels. Podesva's (e.g. 2006, 2007) studies of situational variation among gay professional men have also shown the stylistic use of pitch. Particularly dramatic is the contrast in styles of a young doctor, Heath, who goes from his calm and serious professional persona in the clinic to a 'Gay Diva' persona at a barbecue with friends. At the barbecue, he shows a significantly higher overall f0, as well as a wider f0 range, greater slope and greater and more durative use of falsetto, yielding a 'prissy' effect.

In her ethnography in an arts high school, Pratt (2019) focused on the embodiment of affective style in the student social order. The technical theater (or 'tech') program was distinct from other arts programs in the school with its focus on manual labor, constructing sets, and working with lighting and audio equipment. The tech students were known as handy, tough, and rowdy, adopting a 'badass' and reticent affective style, wearing black clothing and work boots, and carrying work knives indicating their toughness and readiness for manual work. The tech students showed a highly significant backing of /a/ and of /l/ in comparison with their peers in the 'fine arts'. Not only is the sound strikingly different, but the backing in both the vowel and the liquid suggests a generally retracted jaw

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

setting. This kind of embodiment has been invoked in accounts of articulatory setting as underlying stylistic features: Labov (1963:307) proposed that the centralization of the nucleus of /ay/ and /aw/ on Martha's Vineyard was embodied in a close-mouthed articulatory style as an 'object of social affect'. And in a study of California vowels, D'Onofrio, Pratt, & Van Hofwegen (2019) have speculated that an open-jaw setting popularly associated with young California 'laid-back' styles (e.g. surfer dudes and valley girls) is related to a front-to-back compression of the vowel system and a lowering of the front vowels. Using a computer vision algorithm to detect smiling and movement, Voigt, Podesva, & Jurafsky (2014) have found that body movement correlates with increases in pitch, pitch variability, and intensity, and Podesva, Callier, Voigt, & Hilton (2015) have shown that speakers use higher f0 in phrases during which they smile.

Consonant fortition has been less studied, but as a form of intensification, it is a common expressive resource. A striking (unpublished) example is from Eckert's ethnographic work in an elementary school (e.g. Eckert 2011). In this case, a preadolescent girl, who was angry at her boyfriend because he was going to play baseball instead of coming to the sixth grade dance with her, yelled out at him during an otherwise calm conversation with Eckert, "I'm really mad at you right now, [Brian]!". Her anger was expressed both with a higher intonation and intensity on "I'm really mad at you right now", and with a lengthening of the initial consonants in *really*, *mad*, *right*, and *now*. Measuring the relative length of /r/ and the following vowel is problematic because of the difficulty of pinning down the boundary between the two, but the nasals offer an indication of the extent of this emphatic lengthening. Table 1 shows the length of the initial nasals and the following vowels in *now* and *mad* in the angry outburst and in all other occurrences of these two words in the conversation. The initial nasals in the angry version (in bold) constitute over half of the duration of the nasal and vowel sequence, a significantly larger proportion than the normal conversational occurrences.

Other examples of fortition have been observed in ethnographic studies of adults. Podesva (2004) examined the variation in stop bursts as Heath moved from his doctor persona to his 'gay diva' persona. While he released /t/ more frequently in his doctor persona in keeping with his articulate professional style, the burst length and amplitude of released /t/ at the barbecue were significantly greater. Podesva characterized the exaggerated releases at the barbecue as conveying an 'antagonistic attitude' which contributed to the 'prissiness' of the diva persona. Calder (2019) filmed a drag queen in the course of a three-hour makeover for a performance. As part of the South of Market drag scene, this performer resists the ultra-feminine performance of traditional drag, striving for a combative, intimidating, and 'fierce' persona. The center of gravity of /s/ remains extremely high throughout, but as she becomes increasingly female in appearance, the /s/ is transformed by an increase in amplitude into a harsh, unpleasant sound.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

TABLE 1. *Length of initial nasals and following vowels in* now *and* mad *(in seconds). Angry occurrences in bold.*

/N/ NOW	/AW/ NOW	% /N/	/M/ MAD	/AE/ MAD	% /M/
0.055	0.185	0.297	0.046	0.155	0.229
0.018	0.162	0.111	0.03	0.17	0.150
0.04	0.226	0.177	0.089	0.237	0.273
0.03	0.17	0.176	0.08	0.11	0.421
0.057	0.319	0.179			
0.05	0.155	0.323			
0.21	**0.25**	**0.840**	**0.162**	**0.15**	**0.519**

PERCEPTUAL EVIDENCE OF ICONICITY

Bringing together these studies reveals two organizing tendencies of rhematized sociolinguistic variants linked with affect—the frequency code and acoustic fortition. So far, the evidence is in studies of production, and relatively little work examining the affective potential of sociolinguistic variables has explored the perceptual domain. We conducted a series of social evaluation tasks to test whether these tendencies observed in production also emerged in listeners' evaluations of affect, using matched-guise style paradigms with acoustically manipulated stimuli (e.g. Lambert, Hodgson, Gardner, & Fillenbaum 1960; Campbell-Kibler 2007). Specifically, we ask whether a level intonational contour, a vowel with lower F2, and an intensified stop burst induce more negative evaluations of speaker affect than a more dynamic intonational contour, vowel with higher F2, and less intensified stop bursts. In examining these evaluations, we offer a first step in testing whether the iconic associations observed and proposed between these variants and affect are borne out in the ways listeners interpret these variants when framed in an interactional context.

Stimuli

In the following experiments, auditory stimuli were constructed from digitally manipulated tokens of a single-word utterance: the name *Todd*. This name was selected because it includes both a low central vowel and initial voiceless stop, making possible multiple manipulations of the same stimulus. The stimulus was initially produced by one of the authors, a female native speaker of American English from the North Central region of the United States (Labov, Ash, & Boberg 2006), recorded in a sound-attenuated booth. All experimental tokens used as stimuli were digitally manipulated versions of the same baseline token of *Todd*, described below. In all cases, resulting tokens were each embedded in the same dialogic frame, with a different speaker (the other author) producing the 1.3 second utterance "Who are you working with?". The token of *Todd* always followed this frame, with a 0.25 second

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

pause between the two. This posed the critical item, *Todd*, as an answer to an inquiry, with the pragmatic assumption that both speakers were familiar with Todd, and that the second speaker was working with him. At issue for the listener was an evaluation of the second speaker's feelings toward the colleague named Todd.

Note that by embedding the manipulations within this specific name and dialogic frame, as spoken by these particular voices, we necessarily created an interactional context to which listeners will bring their experiences, assumptions, and ideologies. These may have included their evaluations of these specific speakers, ideologies related to the perceived gender of the interactants and the stance object *Todd*, and ideas about why the speakers would be discussing this particular topic and the affect that might be attached to such interactional settings. We chose this context intentionally, as listener evaluations of the affective variables we test here necessarily rely on an interactional context that would evoke a valenced evaluation by the speaker toward some object (here, another individual). Indeed, given the view that every linguistic interpretation is ideologically loaded, listener assumptions about a given speech sample are unavoidable—there is no socially 'neutral' context, as even nonce words produced in completely synthesized voices carry social assumptions (e.g. under what social circumstances might a listener encounter a robotic-sounding voice?), and an evaluation of affect under such circumstances would be quite difficult to interpret.

This particular context may invoke particular affective evaluations regardless of the phonetic variants uttered. It is for just this reason that we keep the context entirely constant across the three manipulations described below, as any CHANGES in listener evaluations between manipulated variants—our measure of interest—can be attributed only to the phonetic manipulations tested. As in any social evaluation task with limited stimuli, the ways in which these particular variants may evoke affective evaluations in other dialogic contexts or when used by other speakers almost certainly will differ (see Campbell-Kibler 2007)—here we test only the potential for these subtle shifts to invoke differences in evaluations of affect, in directions predicted by previous work that has described these variables as rhematized.

Three variable types were manipulated, each examining a different aspect of phonetic detail that has been proposed to have iconic potential in production: pitch contour, vowel backness, and two acoustic dimensions of the stop burst (duration and intensity). The motivation, manipulation of stimuli, and predictions for each of these conditions regarding their potential influences on affective meaning conveyed by the utterance are each described below.

Intonation

In the intonation manipulation, we shifted the F0 contour of the stimulus *Todd* in Praat (Boersma & Weenink 2018) to yield two pitch contours: a falling contour with a high pitch accent, and a level contour with a low pitch accent. Pitch contours

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

TABLE 2. *F0 maxima, minima, and ranges for original and manipulated tokens in the intonation manipulation conditions.*

	ORIGINAL TOKEN (HZ)	FALLING TOKEN	LEVEL TOKEN
Vowel F0 max (Hz)	220	239	189
Vowel F0 min (Hz)	148	146	148
Vowel F0 range (Hz)	72	93	41

are, of course, known to convey referential, pragmatic, and affective meaning, and may also contain iconic potential (e.g. McLemore 1990). In particular, F0 levels and contours have been shown to modulate judgments of affect both independently of referential content and in interaction with it (e.g. Scherer, Ladd, & Silverman 1984; Ladd, Silverman, Tolkmitt, Bergmann, & Scherer 1985). To this end, we manipulated the pitch tracks of the vowel in the baseline token of *Todd* to either follow a more dynamic falling contour, with a higher pitch accent and thus a greater fall than the original token, or a more level contour, with the pitch accent lowered and the F0 nearly level. Details about the resulting tokens' F0 values are shown in Table 2. Both full dialogue tokens were scaled for a mean intensity of 80dB, and both vowel quality and acoustic aspects of the initial stop consonant were retained from the original token.

Participants in one of the intonation conditions heard EITHER the manipulated falling token or the level token; none were exposed to the original baseline token. Given previous work on the emotional or affective indices of these contours (e.g. Jeong & Potts 2016), we predicted that the falling token would yield significantly more positive affective evaluation scores than the level token.

Vowel backness

In the vowel backness manipulation, we used the Akustyk package (Plichta 2013) in Praat (Boersma & Weenink 2018) to manipulate the F2 value of the /ɑ/ vowel in *Todd*. This package allows for resynthesis of specified formant tracks while keeping the token's other formant values, F0, and other acoustic dimensions the same as the original token's. Note that the original baseline F0 trajectory was not altered for either of the vowel backness tokens. The vocalic /ɑ/ portion of the target word *Todd* was spliced out of the /t_d/ frame, and resynthesis was specified to set F2, for any given timestep throughout the vowel, at either 200Hz less than the original for the backed token, or 200Hz greater than the original for the fronted token. As resynthesis adjusts the formant values along many timesteps across the vowel, this allowed the vowel to retain its original formant dynamics while shifting the entirety of the trajectory of the second formant up or down 200Hz. F2 measurements at the vowel midpoint are provided in Table 3.

The manipulated vowels were then spliced back into the original /t_d/ frame and embedded into the dialogue as described above. Both full dialogue tokens

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

TABLE 3. *F2 midpoint measurements for original and manipulated tokens in the vowel backness condition.*

	ORIGINAL TOKEN (HZ)	BACKED TOKEN (HZ)	FRONTED TOKEN (HZ)
F2 at midpoint	1499	1294	1683

were scaled for a mean intensity of 80 dB. Again, these participants were never exposed to the original baseline token; they heard either the backed or fronted token. Neither the intonational contour nor acoustic aspects of /t/ were manipulated from the original token in this condition. Given findings in production relating backed F2 to negative affect (Eckert 2008; Geenberg 2014), we predicted that the fronted token would yield significantly more positive affective evaluation scores than the backed token.

Stop burst duration and intensity

The final aspect of the target token that was manipulated was the burst of the voiceless stop /t/ in *Todd*. As mentioned above, Podesva (2006) showed that intensification of social meaning can be achieved via the lengthening of stop bursts, a form of fortition that increases acoustic prominence. Another dimension of acoustic prominence, intensity, may analogically intensify the social meaning of a given segment as well. Intensification via either increased length or amplitude of bursts can be thought to not only iconically intensify the social meaning linked with the segment itself, it can also be rhematized as part of a precise, pointed, or forceful style, associable with anger, frustration, or disgust. To this end, our final manipulation of the *Todd* stimuli focused on the burst in the word-initial voiceless stop /t/. We manipulated both duration and intensity to examine their individual contributions to perceived affect as well as the ways that these dimensions may work together to achieve intensification.

Manipulation of the stop burst was performed in Praat (Boersma & Weenink 2018). As in the other stimuli, the entire dialogue token was scaled at a mean intensity of 80 dB. Then, the stop burst was spliced out of the *Todd* frame. Original duration (ms) and intensity (dB) values of the burst are shown in Table 4. For the duration manipulation, the burst was either lengthened or shortened in Praat. A series of manipulations were tested, and the authors determined the largest audible difference between tokens that both still sounded like natural productions. The resulting 'long' token doubled the duration of the original to 64ms, while the 'short' token was 25 ms, 75% of the original duration. The 'long' and 'short' bursts were then each manipulated for mean intensity. Intensity of the burst was measured in relation to the mean intensity of the following vowel (82 dB). Once again, a series of intensity manipulations were attempted in Praat. While we attempted decreased intensity manipulations, these sounded unnatural. Our 'lower intensity' token

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

TABLE 4. *Burst duration and intensity measurements for original and manipulated tokens in the stop burst conditions.*

	Original token	Lengthened token	Shortened token
Burst duration (ms)	32	64	25
	Following vowel	Increased stop burst intensity	Original stop burst intensity
Mean intensity (dB)	82	75	65

therefore retained the original token's mean burst intensity of 65dB, 80% of the intensity of the following vowel. The largest possible difference that still sounded plausibly natural was a 10 dB increase in mean burst intensity to 75 dB, 90% of the following vowel.

The two manipulated burst durations (long and short) were crossed with the two intensities (higher and lower) for a total of four burst tokens (long-high, long-low, short-high, short-low). Note that rather than treating these as two separate manipulations, we cross these dimensions given that they entail manipulations of the same segment. We also wished to assess whether the two dimensions of fortition could operate in concert with one another, whether cumulatively or separately. Vowel backness and intonational contour were retained from the original token, and no other aspects of the tokens were manipulated. Again, participants only heard one of these four tokens. We predicted that increased duration and increased intensity would cue more negative ratings of affect as compared to the shorter or lower intensity tokens. We were also particularly interested in the ways that these two dimensions of intensification may behave in concert with one another to cue negative affect. Social meanings of linguistic styles are necessarily derived by constellations of linguistic features (Eckert 2008), and perceptual work has shown that combinations of features contribute differently to social evaluations than single feature manipulations (Levon 2007; Campbell-Kibler 2011). Here, we examine whether similar findings can be observed for two acoustic dimensions of a single segment, both predicted to affect intensification in ways that cue negative affect.

Design and procedure

Across the three stimulus types, the experiment tested a total of eight conditions: two conditions for pitch contour (falling and level), two for vowel backness (back and front), and four for stop burst (long vs. short duration crossed with high vs. low intensity). Every condition of the study was balanced between subjects, such that any given individual listener heard the auditory dialogue only once, with only one of the eight tokens of the target item *Todd*.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

Listeners were recruited and compensated on Mechanical Turk, and they were directed to participate in the main task via an external link to a Qualtrics survey. Only participants whose worker ID identified them as from the United States were recruited. Listeners were first presented with a consent form, followed by a sound check that asked them to type an auditory word that was played for them, spoken in a different voice than those used in the main task. This ensured that the participants were able to hear audio for the experiment.

All listeners were then presented with the auditory stimulus dialogue, again, with the token of *Todd* differing between subjects. The stimulus was accompanied by the following instructions.

> Please listen to this brief recording of a dialogue between two people:
> A: Who are you working with?
> B: [person's name]

The accompanying transcript was intended to focus the listeners' attention on Speaker B's utterance, as well as to label the speakers as Speaker A and B for reference in the following questions. Listeners clicked a 'play' button to play the dialogue, then they were asked three questions. First, they were asked 'Who is speaker B working with?'. This was to ensure that the participant accurately heard the utterance as 'Todd'. Second, speakers were asked, 'How does speaker B feel about working with this person?', accompanied by a five-point Likert scale. The points were each labeled from 1–5: very negative, somewhat negative, neither negative nor positive, somewhat positive, very positive. This wording was chosen to ensure that the participant was evaluating the affect speaker B was displaying TOWARD TODD (and not toward speaker A, toward some other object, or generally). This offered a level of consistency in what was being evaluated across listeners, focused the question on the lexical item that was manipulated, and ensured that the dialogic context was taken into account in evaluation by all listeners. Third, they were asked to optionally provide additional comments or impressions about the dialogue. While a numerical scale of affect is quite literally a one-dimensional measure, we follow previous work using social evaluations in sociolinguistics that utilize attribute scales, in order to quantify the subtle differences in evaluations that may have arisen. We couple this with qualitative assessments of more open-ended comments.

Following the evaluation task, participants were asked to self-report their age, gender, what US state they grew up in and currently live in, and their native languages. They also reported whether they had known hearing disorders, what type of listening device was used in the task (headphones, earbuds, or computer speakers), and they rated the noise level of their environment during the task on a ten-point scale from extremely quiet to extremely noisy.

Participants

All listeners participated in the experiments online, with recruiting and compensation via Amazon's Mechanical Turk (see Schnoebelen & Kuperman 2010;

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

Buhrmester, Kwang, & Gosling 2011; Eskénazi, Levow, Meng, Parent, & Suendermann 2013; and Yu & Lee 2014 for discussions of the benefits and validity of use of Mechanical Turk for linguistics and speech processing research). Fifty participants were recruited per condition, for a total of 400 participants across the eight conditions. Only listeners who self-reported that they grew up in and currently lived in the US and reported a native language as English were retained for analysis. Because of the nature of an online listening task, listeners were also eliminated if they reported any known hearing disorder, or if they reported that their environment was at a noise level above five on the ten-point scale. Participants were also removed if they did not accurately report speaker B to be saying 'Todd'. Total numbers of participants submitted to final analysis are shown by condition in Table 5.

Analysis

Results of the affective scale ratings (very positive to very negative) were assessed within variable type (intonation, vowel backness, or burst). The five-point affective rating scale was converted to a numeric vector and assessed as the continuous dependent variable in a series of linear regression models in R (R Core Team 2013). Note that since this study was conducted entirely between subjects, and each participant provided only a single data point for affective ratings, typical random effects of participant and item did not apply in the dataset, and thus a simple linear regression model was used.

One linear regression model was fit to the rating set for each variable type, for a total of three models. In all cases, the stimulus condition was included as a categorical predictor with either two (intonation, backness) or four (burst) levels. Participant age, gender, and dialect region according to the Atlas of North American English (Labov et al. 2006) were tested as predictors in all three models, but they did not prove to be a significant predictor nor improve model fit, and they were not included in final models.

RESULTS

Overall, participants tended to rate the dialogue on the 'negative' end of the scale regardless of the stimulus manipulations, with an overall mean rating of 2.5 (between 'somewhat negative' and 'neither negative nor positive'). Each of the variable manipulations yielded small but statistically significant differences in how listeners perceived Speaker B to feel about 'Todd', with means and 95% confidence intervals shown in Figure 1.

The F0 contour manipulation showed the largest difference in ratings, in the expected direction: the level contour was rated significantly lower on the scale than the falling contour (Beta = -0.44, SE = 0.15; T = -2.92; $p = 0.005$). There tended to be agreement overall, with 86% of the respondents (N = 67) in these manipulation conditions answering either 'somewhat negative' or 'neither negative or positive',

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

TABLE 5. *Number of participants in each condition.*

CONDITION	N PARTICIPANTS IN FINAL ANALYSIS
Intonation: Falling	43
Intonation: Level	40
Backness: Front	39
Backness: Back	41
Stops: Long + High Intensity	46
Stops: Long + Low Intensity	49
Stops: Short + High Intensity	48
Stops: Short + Low Intensity	50
Total N	356

9% (N = 8) responding 'somewhat positive', 4% (N = 3) responding 'very negative', and none responding 'very positive'. Comments on the voice in the level condition described Speaker B as 'unhappy' and 'irritated' by multiple participants, with one describing the utterance as "unpleasant for her to even say the name. 'Todd' sounded like 'pond scum'". In the falling condition, most participants described the utterance as 'neither positive or negative', 'neutral', or 'indifferent', though some described the voice as "annoyed". This condition was included to serve as a baseline test for the manipulation and scale—within this dialogue, affect toward 'Todd' was assessed differently depending on acoustic measures.

Vowel backness (F2) also significantly shifted ratings, again in the direction predicted by the frequency code and shown in previous sociophonetic work. The backed token was rated as significantly more negative than the fronted token (Beta = 0.35, SE = 0.17, T = 2.08, $p = 0.041$). Again, there was a good deal of consensus in the distribution of responses, with 85% of respondents in these condition (N = 68) responding either 'somewhat negatively' or 'neither negative or positive', 11% (N = 9) responding 'somewhat positive', 3% (N = 2) responding 'very negatively', and 1% (N = 1) responding 'very positively'. In the backed condition, Speaker B was described as "annoyed", "not enthused", "angry", and "exasperated". While "annoyed" and "frustrated" also came up in comments in response to the fronted condition, more commenters described the voice as "neutral" and "matter-of-fact".

The differences were the smallest in the burst manipulation conditions. Again, the distribution was concentrated around 'somewhat negative' or 'neither negative or positive', with 89% of respondents (N = 172) answering one of these, 8% (N = 15) answering 'somewhat positive', 3% (N = 5) 'very negative', and 0.5% (1) 'very positive'. It was not the case that a given dimension (duration or intensity) shifted ratings overall. However, as predicted, the most negative ratings were assigned to the token with both forms of intensification: the long burst duration and

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

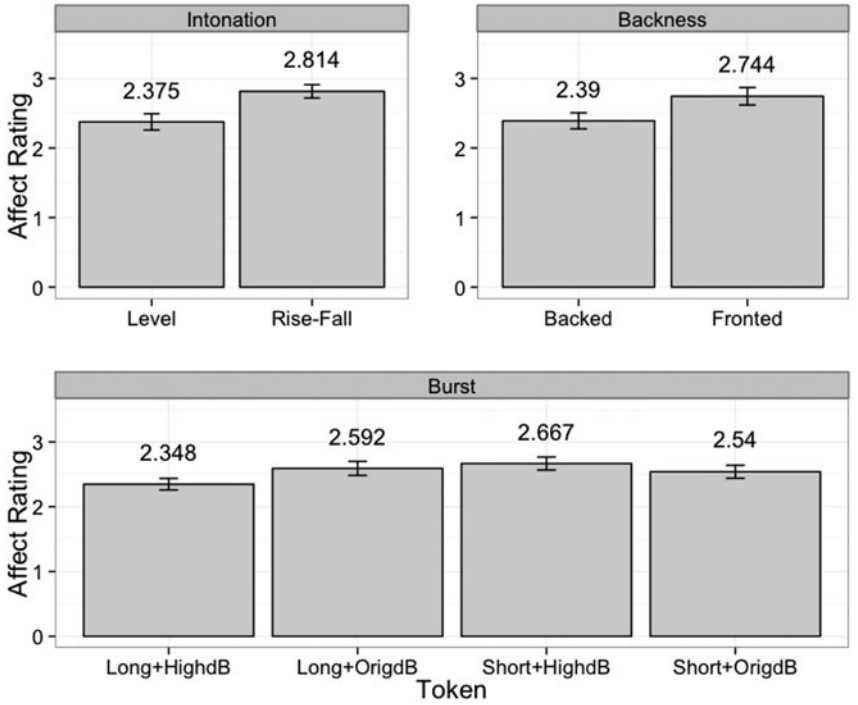


FIGURE 1. Perceptions of Speaker B's affect by intonation, vowel backness, and burst length.

high burst intensity. This stimulus showed significantly lower ratings from the token with short duration and high intensity (Beta = 0.32, SE = 0.14, T = 2.22, $p = 0.027$), though it was not significantly different from the long token with original intensity (Beta = 0.24, SE = 0.14, T = 1.71, $p = 0.09$), nor from the short token with original intensity (Beta = 0.19, SE = 0.14, T = 1.35, $p = 0.18$).

In the most negatively rated burst condition (longer duration, higher intensity), participants frequently described Speaker B as "annoyed", "angry/mad", "disgusted", and "irritated". One comment explains, "It's obvious from the tone of voice that the person is annoyed. She could be annoyed with the person asking the question but I thought it more likely that she isn't all that impressed with Todd". In the highest rated burst condition (shorter duration, higher intensity), participants provided more comments about the "neutral" or "indifferent" affect of the second speaker. The two middle conditions showed a mix between "neutral" comments and "annoyed" comments.

The significant difference in ratings according to burst duration arose between the two high intensity tokens. However, intensity did not appear to consistently affect ratings on its own—the most positively and the most negatively rated

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

tokens were both those with a high intensity burst. This follows Podesva's (2006) finding that burst duration, but not peak frequency, corresponded with context- and discourse-specific expressions of a 'prissy' persona in Heath's stylistic practice. However, we note that this effect appears to be even more nuanced—the burst duration difference only arises in the high intensity tokens, but no difference in duration emerged between the two lower intensity tokens. For this particular dialogue, a reading of negative affect requires the coalescence of cues to acoustic intensification —a longer AND a louder burst. It may be the case that lengthening without increased loudness, or loudness without an increase in duration, may cue other social meanings or be used stylistically to different ends untapped by our single affective rating scale.

Together, the *Todd* tasks reveal that aspects of segmental variation—vowel quality, and stop burst and duration—correspond to listener evaluations of affect in ways that reflect iconic processes of the frequency code and fortition to similar degrees as more conventionalized intonational contours. As shown in the case of consonant fortition, intensification of meaning can occur iconically through the phonetic intensification of a particular segment—longer and louder bursts combine to yield stronger affective meanings from the same phonological segment.

CONCLUSIONS

Until quite recently, affect has been ignored in sociolinguistics, relegated to 'individual' rather than 'social' dynamics. But affect is eminently social, based in human relations and interaction, and built into distinctions of class, race, gender, and age (Besnier 1990). Irvine (1990) offers a particularly dramatic example from a Wolof community, where affective display is central to the caste system, and encoded at every level of the grammar. And the encoding is eminently iconic, from speech rate to pitch variability to noun class assignments.

The manipulations of phonetic cues in our experiments are quite subtle; nonetheless, they consistently shift listener evaluations of speaker affect. This suggests that observations of iconicity in production studies are not the result of overinterpretation, but evidence of meaningful processes. Iconicity is not natural, but an engagement with things we have come to see as natural. The social distinctions that attract the indexicality of variation are sufficiently important to our beliefs that we strive to see them as inevitable. Following Jakobson (1965:25, citing Benveniste 1939), one might say that it is not the link between sound and meaning that is natural, but the need to make the link:

> only for a detached, alien onlooker is the bond between the signans and signatum a mere contingence, whereas for the native user of the same language this relation changes into a necessity.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

REFERENCES

Acton, Eric (2019). Pragmatics and the social life of the English definite article. *Language* 95:37–65.

———, & Christopher Potts (2014). That straight talk: Sarah Palin and the sociolinguistics of demonstratives. *Journal of Sociolinguistics* 18:3–31.

Agha, Asif (2003). The social life of a cultural value. *Language and Communication* 23:231–73.

Austin, John L. (1962). *How to do things with words*. Oxford: Oxford University Press.

Bekker, Ian (2007). Fronted /s/ in General White South African English. *Language Matters* 38:46–74.

———, & Erez Levon (2017). The embedded indexical value of (s)-fronting in Afrikaans and South African English. *Linguistics* 55:1109–39.

Benveniste, Emile (1939). Nature du signe linguistique. *Acta Linguistica* 1:23–29.

Besnier, Niko (1990). Language and affect. *Annual Review of Anthropology* 19:419–51.

Bloomfield, Leonard (1933). *Language*. New York: Henry Holt.

Boersma, Paul, & David Weenink (2018). *Praat: Doing phonetics by computer*, version 6.0.16. Online: www.praat.org; accessed 1 September 2018.

Bolinger, Dwight (1950). Rhyme, assonance, and morpheme analysis. *Word* 6:117–36.

——— (1985). The inherent iconism of intonation. In John Haiman (ed.), *Iconicity in syntax: Proceedings of a symposium on iconicity in syntax, Stanford, June 24–26, 1983*, 97–108. Amsterdam: John Benjamins.

——— (1989). *Intonation and its uses*. Stanford, CA: Stanford University Press.

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* 16:3–5.

Butler, Judith (1993). *Bodies that matter*. Abingdon: Routledge.

Calder, Jeremy (2019). The fierceness of fronted (s): Linguistic rhematization through visual transformation. *Language in Society* 48:31–64.

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

Campbell-Kibler, Kathryn (2011). Intersecting variables and perceived sexual orientation in men. *American Speech* 86:52–68.

Casasanto, Daniel, & Katinka Dijkstra (2010). Motor action and emotional memory. *Cognition* 115:179–85.

Clark, Herbert H. (2006). Context and common ground. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., vol. 3, 105–108. Oxford: Elsevier.

de Saussure, Ferdinand (1916). *Cours de linguistique générale*. Paris: Payot.

D'Onofrio, Annette (2014). Phonetic detail and dimensionality in sound-shape correspondences: Refining the bouba-kiki paradigm. *Language and Speech* 57:367–93.

——— (2015). Persona-based information shapes linguistic perception: Valley girls and California vowels. *Journal of Sociolinguistics* 19:241–56.

———; Teresa Pratt; & Janneke Van Hofwegen (2019). Compression in the California Vowel Shift: Tracking generational sound change in California's Central Valley. *Language Variation and Change* 31(2):193–217.

Dingemanse, Mark (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics* 3:1–30. doi: http://doi.org/10.5334/gjgl.444.

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

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

——— (2011). Where does the social stop? In Frans Gregersen, Jeffrey K. Parrott, & Pia Quist (eds.), *Language variation: European perspectives III*, 13–30. Amsterdam: John Benjamins.

——— (2019a). The individual in the semiotic landscape. *Glossa: A Journal of General Linguistics* 4. doi: http://doi.org/10.5334/gjgl.640.

——— (2019b). The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language* 95:751–76.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

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

Flipsen Jr., Peter; Lawrence Shriberg; Gary Weismer; Heather Karlsson; & Jane McSweeny (1999). Acoustic characteristics of (s) in adolescents. *Journal of Speech, Language and Hearing Research* 42:663–77.

Fónagy, Ivan (1983). *La vive voix: Essais de psycho-phonétique*. Paris: Payot.

Gabrielsson, Alf, & Erik Lindström (2010). The role of structure in the musical expression of emotions. In Patril N. Juslin & John A. Sloboda (eds.), *Handbook of music and emotion: Theory, research, applications*, 367–400. Oxford: Oxford University Press.

Gal, Susan (2013). Tastes of talk: Qualia and the moral flavor of signs. *Anthropological Theory* 13:31–48.

Gal, Susan, & Judith T. Irvine (2019). *Signs of difference*. Cambridge: Cambridge University Press.

Geenberg, Kate (2014). Sound symbolism in Adult Baby Talk (ABT): The role of the frequency code in the construction of social meaning. Paper presented at NWAV 43, Chicago, IL.

Gordon, Matthew, & Jeffrey Heath (1998). Sex, sound, symbolism, and sociolinguistics. *Current Anthropology* 39:421–49.

Gussenhoven, Carlos (2016). Foundations of intonational meaning: Anatomical and physiological factors. *Topics in Cognitive Science* 8:425–34.

Haeri, Niloofar (1996). 'Why do women do this?': Sex and gender differences in speech. In Gregory R. Guy, Crawford Feagin, Deborah Schiffrin, & John Baugh (eds.), *Towards a social science of language*, 101–14. Amsterdam: John Benjamins.

Haiman, John (1980). The iconicity of grammar. *Language* 56:5–40.

Holmes, Elliott, & Erez Levon (2017). The substance of style: Gender, social class and interactional stance in (s)-fronting in southeast England. *Linguistics* 55:1045–72.

Hubbard, Timothy L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American Journal of Psychology* 109:219–38.

Irvine, Judith (1990). Registering affect: Heteroglossia in the linguistic expression of emotion. In Catherine A. Lutz & Lila Abu-Lughod (eds.), *Language and the politics of emotion*, 121–61. New York: Cambridge University Press.

——— (2001). Style as distinctiveness: The culture and ideology of linguistic differentiation. In Penelope Eckert & John R. Rickford (eds.), *Stylistic variation in language*, 21–43. Cambridge: Cambridge University Press.

———, & Susan Gal (2000). Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, politics, and identities*, 35–83. Santa Fe, NM: SAR Press.

Jakobson, Roman (1960). Closing statement: Linguistics and poetics. In Thomas A. Sebeok (ed.), *Style in language*, 350–77. New York: MIT Press and John Wiley & Sons.

——— (1965). Quest for the essence of language. *Diogenes* 13:21–37.

———, & Linda R. Waugh (2002). *The sound shape of language*. 3rd edn. Berlin: Mouton de Gruyter.

Jeong, Sunwoo (2016a). Conventions for affective meanings: Non-canonical terminal contours in English polar interrogatives. *Speech Prosody*, 907–11.

——— (2018). Intonation and sentence type conventions: Two types of rising declaratives. *Journal of Semantics* 35:305–56.

———, & Christopher Potts (2016). Intonational sentence-type conventions for perlocutionary effects: An experimental investigation. *Semantics and Linguistic Theory* 26 (SALT 26):1–22.

Joseph, John (2015). Iconicity in Saussure's linguistic work, and why it does not contradict the arbitrariness of the sign. *Historiographia Linguistica* 42:85–105.

Köhler, Wolfgang (1929). *Gestalt psychology*. New York: Liveright.

Labov, William (1963). The social motivation of a sound change. *Word* 18:273–309.

——— (1972). On the mechanism of linguistic change. In *Sociolinguistic patterns*, 160–82. Philadelphia: University of Pennsylvania Press.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

———; Sharon Ash; & Charles Boberg (2006). *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.
Ladd, D. Robert (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
———; Kim Silverman; Frank Tolkmitt; Günther Bergmann; & Klaus Scherer (1985). Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *Journal of the Acoustical Society of America* 78:435–44.
Lambert, Wallace; Richard Hodgson; Robert Gardner; & Samuel Fillenbaum (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology* 60:44–51.
Levon, Erez (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society* 36:533–54.
———; Marie Maegaard; & Nicolai Pharao (2017). *The sociophonetics of /s/*. Special issue of *Linguistics* 55(5).
Li, Fangfang (2017). The development of gender-specific patterns in the production of voiceless sibilant fricatives in Mandarin Chinese. *Linguistics* 55:1021–44.
Maurer, Daphne; Thanujeni Pathman; & Catherine J. Mondloch (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science* 9:316–22.
McLemore, Cynthia (1990). The interpretation of L*H in English. In Cynthia McLemore (ed.), *Texas linguistic forum 32: Discourse*, 127–47. Austin: University of Texas Department of Linguistics and Center for Cognitive Science.
Morton, Eugene S. (1994). Sound symbolism and its role in non-human bertebrate communication. In Leanne Hinton, Johanna Nichols, & John J. Ohala (eds.), *Sound symbolism*, 348–65. Cambridge: Cambridge University Press.
Ohala, John (1994). The biological bases of sound symbolism. In Leanne Hinton, Johanna Nichols, & John J. Ohala (eds.), *Sound symbolism*, 222–36. Cambridge: Cambridge University Press.
Peirce, Charles S. (1868). On a new list of categories. *Proceedings of the American Academy of Arts and Science* 7:287–98.
Pennycook, Alastair (1985). Actions speak louder than words: Paralanguage, communication, and education. *TESOL Quarterly* 19:259–82.
Plichta, Bartlomiej (2013). Akustyk: Speech analysis and synthesis plug-in for Praat. Online: http://github.com/akustyk.
Podesva, Robert (2004). On constructing social meaning with stop release bursts. Paper presented at Sociolinguistics Symposium 15, Newcastle upon Tyne.
——— (2006). Intonational variation and social meaning: Categorical and phonetic aspects. *University of Pennsylvania Working Papers in Linguistics* 12:189–202.
——— (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11:478–504.
———; Patrick Callier; Rob Voigt; & Katherine Hilton (2015). The voice embodied: Bringing the quantitative analysis of body movement into the study of phonation. Poster presented NWAV 44, Toronto.
———, & Janneke Van Hofwegen (2016). (s)exuality in small-town California: Gender normativity and the acoustic realization of (s). In Erez Levon & Ronald Beline Mendes (eds.), *Language, sexuality and power: Studies in interactional sociolinguistics*, 168–88. Oxford: Oxford University Press.
Potts, Christopher (2007). The expressive dimension. *Theoretical Linguistics* 33:165–97.
Pratt, Teresa (2019). Embodying 'tech': Phonetic variation, articulatory retraction, and social meaning. *Journal of Sociolinguistics* 23:1–22.
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/.
Ramachandran, V. S., & E. M. Hubbard (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies* 18:3–34.
Sapir, Edward (1929). A study in phonetic symbolism. *Journal of Experimental Psychology* 12:225–39.
Scherer, Klaus R.; D. Robert Ladd; & Kim E. A. Silverman. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America* 76:1346–56.

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press

Schnoebelen, Tyler, & Victor Kuperman (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija* 43:441–64.
Silverstein, Michael (1976). Shifters, linguistic categories, and cultural description. In Keith H. Basso & Henry A. Selby (eds.), *Meaning in anthropology*, 11–55. Albuquerque: University of New Mexico Press.
——— (1994). Relative motivation in denotational and indexical sound symbolism of Wasco-Wishram Chinookan. In Leanne Hinton, Johanna Nichols, & John J. Ohala (eds.), *Sound symbolism*, 40–60. Cambridge: Cambridge University Press.
——— (2003). Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23:193–229.
Stuart-Smith, Jane (2007). Empirical evidence for gendered speech production: (s) in Glaswegian. In Jennifer Cole & Jose I. Hualde (eds.), *Laboratory phonology 9*, 65–86. New York: Mouton de Gruyter.
Tagliamonte, Sali, & Alexandra D'Arcy (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85:58–108.
Trager, George L. (1958). Paralanguage: A first approximation. *Studies in Linguistics* 13:1–12.
Voigt, Robert; Robert J. Podesva; & Dan Jurafsky (2014). Speaker movement correlates with prosodic indicators of engagement. *Proceedings of Speech Prosody* 7. Online: https://nlp.stanford.edu/pubs/speechprosody_voigt.pdf.
Woodworth, Nancy L. (1991). Sound symbolism in proximal and distal forms. *Linguistics* 29:273–99.
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.
——— (2008). Rhotacization and the 'Beijing smooth operator': The social meaning of a linguistic variable. *Journal of Sociolinguistics* 12:201–22.
——— (2018). *Language and social change in China: Undoing commonness through Cosmopolitan Mandarin*. New York: Routledge.

(Received 26 February 2020; revision received 09 September 2020;
accepted 17 September 2020; final revision received 25 September 2020)

Address for correspondence:
Penelope Eckert
Department of Linguistics
Stanford University
Stanford, CA 94305, USA
eckert@stanford.edu

https://doi.org/10.1017/S0047404520000871 Published online by Cambridge University Press