Journal of Sociolinguistics

WILEY

Journal of Sociolinguistics

DIALOGUE: COMMENTARY

Cross-Journal Symposium

In-Group/Out-Group Dynamics, Contrast, and the Listening Subject in Sociolinguistic Perception

Lacey Wade

University of Kansas, Lawrence, Kansas, USA

Correspondence: Lacey Wade (laceywade@ku.edu)

Received: 24 September 2025 | **Revised:** 25 September 2025 | **Accepted:** 26 September 2025

ABSTRACT

Research in linguistic perception has shown that social knowledge shapes how speech is processed, with listeners' social biases influencing their interpretation of the speech signal. Such findings are inherently in step with the concept of *listening practices* in linguistic anthropology. Miyako Inoue's foundational work on the "listening subject" highlights how meaning emerges through listeners' ideological positions, demonstrated through an analysis of Japanese schoolgirl speech. I argue that sociolinguistic perception work has much to be gained from integrating a listening subject framework, foregrounding the social conditions enabling perception and emphasizing interpretive agency. In this commentary, I focus on how the listening subject is constituted through contrast with the object of speech perception and explore how this dynamic intersects with ongoing work on in-group/out-group identity and the role of experience in shaping sociolinguistic perception, specifically drawing on my research on US regional dialect perception. Finally, I discuss how experimental sociolinguistics and public engagement efforts might be enhanced by critically examining the listening subject positions of research participants.

1 | Introduction

Work in linguistic perception has demonstrated that social knowledge influences linguistic processing, and vice versa. A growing body of work shows that the same acoustic signal is interpreted differently depending on the perceived social characteristics of the speaker (e.g., D'Onofrio 2015, 2018; Casasanto 2008; Koops et al. 2008; Hay et al. 2006). In some cases, listeners' own biases can even *override* the acoustic input entirely (e.g., Niedzielski 1999; Wade et al. 2023), leading them to report hearing features that are not actually present in the signal speech signal. Linguistic features are also evaluated along socially meaningful dimensions (e.g., Campbell-Kibler 2007; Labov et al. 2011). All of these findings underscore the social context in which listening takes place.

Work on *listening practices* in linguistic anthropology is inherently in step with such findings in that it positions listening as a social practice. Inoue (2003) illustrates this in her analysis of 19th-century Japanese schoolgirl speech, arguing that its indexical link to the "schoolgirl" figure emerges not from the speakers' own construction of meaning, but through its interpretation, filtered through the ears of the elite male intellectuals who overheard it. These men, occupying what Inoue terms the "listening subject," played an active role in meaning-making—not merely hearing but interpreting speech in ways shaped by their own social anxieties. They viewed schoolgirl speech as emblematic of unsettling social change, projecting their anxieties about modernization onto the voices of young women, rendering the listening subject—not the speaking schoolgirl—the primary site of semiotic production. This account thus reframes indexicality as

This commentary is part of a Cross-Journal Symposium on Listening Practices and Linguistic Perception in which early career scholars engage with Inoue (2003) and six additional articles published in the *Journal of Linguistic Anthropology* and the *Journal of Sociolinguistics*. All papers are assembled in a virtual issue, available at https://onlinelibrary.wiley.com/doi/toc/10.1002/(ISSN)9999-0009.Listening-Practices-and-Linguistic-Perception.

something conferred by listeners, foregrounding the ideological investments that underpin the social act of listening.

The *listening subject* has obvious connections with work in sociolinguistic perception. For one, framing the listener as the source of indexical meaning is in line with empirical research on *sociolinguistic coherence*. This line of work has found that styles and lects—such as "formal speech" or the "Jewish ethnolinguistic repertoire" (Benor 2011)—are not necessarily internally "coherent," such that all variants associated with a given lect to co-vary or display an implicational relationship. Rather, speakers can variably draw on features of a given lect to construct particular social identities. The coherence of the lect therefore does not necessarily reside in the empirical correlation of its features but in the *perception* of the features as coherent. Indeed, recent studies have found greater coherence in perception than in production (Bennis and Hinskens 2022; Gregersen and Pharao 2016; Wong and Babel 2017; Moore and Montgomery 2018), underscoring the point that lects and styles operate as real social objects not by virtue of statistical consistency but because listeners perceive and impose coherence upon them.

This interpretive agency extends to the moment of perception itself. Experimental work has shown that, even when the physical properties of a speech signal remain constant, the listener transforms them by way of interpreting them (e.g., D'Onofrio 2018; Staum-Casasanto 2010; Hay et al. 2006). Edwards' (2018) discussion of Kockelman's (2010) "channels" further highlights the interpretive agency of the listening position: "Semiotic agents don't just relay signs—they receive them, develop them, and direct them toward other semiotic agents…each channel, each network, has properties, capacities, and vulnerabilities of its own, which means that signals can be intercepted, tampered with, dampened, or lost entirely…indeed, what goes into a channel is rarely the same as what comes out" (275).

Moreover, the layered influence of both hegemonic cultural ideologies and individual positionalities also highlights the relevance of the listening subject to sociolinguistic perception. For instance, D'Onofrio (2019) argues that listeners conferred greater attentional weight to speech presented in an East Asian guise (compared to a White guise), driven by widespread ideological assumptions about accentedness and linguistic difference. At the same time, Wong and Babel (2017) found that ethnic identification based on voice was highest for social groups that matched those of the individual listeners, but also for ideologically favored White voices. Both studies highlight how listening is a product of both the *individual listeners' positionalities* as well as *broader ideological structures* that frame some voices as marked and others as unremarkable.

Having established some foundational overlap between the concept of the listening subject and empirical findings in sociolinguistic perception, the commentary that follows explores more targeted points of connection. I specifically focus on how the listening subject is constituted through *contrast* with the object of speech perception, and how this dynamic intersects with ongoing work on in-group/out-group identity and the role of exposure and experience in shaping sociolinguistic perception, drawing from my own work on the perception of US regional dialects.

2 | The Listening Subject and Contrasting Identities in Sociolinguistic Perception

Central to the concept of the *listening subject* is its construction *in opposition to* the speaking object. In Inoue's (2003) foundational account, stark contrasts are drawn between the "schoolgirl" and the "male intellectual" in age, gender, and position within a modernizing society. Male intellectuals were positioned as "overhearers," exhibiting a type of "linguistic voyeurism," rather than as true interlocutors in conversation with the object of their listening—highlighting their outsider perspective. The authority of their interpretation stemmed not from proximity or participation but from *alterity*: They were older, male, elite, and ideologically dislocated from the youthful, feminized voice they rendered meaningful.

Yet, as Reyes (2017) demonstrates, contrasts need not be stark to be socially productive. In work on the *conyo* elite in the Philippines, Reyes argues that the listening subject—here the middle-class elite—can claim an intimate familiarity while still maintaining an evaluative distance. Through bifurcation of the elite into the frivolous *conyo* and the sensibly moral "middle-class"—Reyes reveals how listeners, even while claiming authority through proximity, ultimately cast the object of their evaluation in opposition to their own social position.

Sociolinguistic perception is often similarly structured through contrast: between accents and the imagined "neutral," local and non-local, in-group and out-group. In our ongoing work on speech perception in Kansas, for example, Kansans often frame their own speech not by what it *is* but by what it *is not*. One participant remarked, "We're Midwestern, so I don't perceive us as having an accent… It's not valley girl, and it's not hard Southern drawl either. It's just kind of middle." Another stated about Kansas speech, "I can only describe it for what it is not." These responses illustrate that the default status of speech ideologically positioned as "neutral" only becomes meaningful in contrast to socially marked others.

And indeed, in-group-out-group distinctions serve not only to define and delineate the language varieties involved—imbuing meaning in one through contrast with the other—but they also shape the way language is perceived. For instance, when asked about how "accented" and "how likely to be from Kansas" various talkers with a range of US dialects were (Wade et al., forthcoming), Kansans showed strong negative relationships between these two scales, choosing voices perceived as the least accented as most likely to be from Kansas. Non-Kansans showed the opposite pattern, choosing talkers judged as most Southern-accented as most Kansas-sounding. This asymmetry highlights how both individual position as in-group or out-group member, as well as broader cultural ideologies placing certain voices as "neutral" or "unaccented" work together to influence listening.

Such perceptual contrasts are closely tied to experience—both in terms of amount and type. In Wade et al. (2023), we examined how listeners' familiarity with Southern US English shaped phonetic convergence in response to sociolinguistic expectations. Non-Southerners with little direct experience converged toward a stereotypically Southern /ay/ vowel after simply being told the

speaker was Southern—even when the speaker was, in fact, from northeastern Ohio. In contrast, Southern listeners, with greater exposure to variation within Southern speech, converged only when they heard actual acoustic cues to Southernness, regardless of what they were told. In other words, inexperienced listeners relied on indexical stereotypes ("Southerners use monophthongal /ay/"), while experienced listeners drew from more nuanced mental representations built through lived exposure.

These findings echo results from Drager and Kirtley (2016), who found that listeners without military experience were more likely to associate Southern-accented speech with military identity—a pattern absent among listeners with military backgrounds. Such differences align with predictions from social psychological research on the *Outgroup Homogeneity Effect* (Park and Rothbart 1982), which holds that out-group members tend to be perceived as more uniform, while in-group members are represented with more nuance and heterogeneity. For out-group listeners in particular, lack of direct experience leads to reliance on stereotyped, often media-generated, representations—what Inoue describes as "fragmented and decontextualized citations" that stand in for knowledge. Similar to the way that stereotypes involve simplistic mental representations that serve as cognitive shortcuts, out-group experience garnered from exaggerated media portrayals and incomplete metalinguistic commentary allow for speech that is "accessible in mediated, imagined, and consumable forms" (Inoue 2003, 160). It is important to note that experience is inherently related to dominance, as Wong and Babel (2017) point out, as listeners often have more linguistic exposure to culturally dominant groups, and familiarity and ideologies may serve to inform and reinforce one another.

While the *listening subject* is generally positioned as emerging through contrast, it is also worth considering how the listening subject may be constituted *within* communities and from an *insider* perspective. This question is relevant to work in sociolinguistic perception, which has found that in-group perceptions tend to be more veridical and nuanced than those of out-group members. One answer lies in acknowledging that any so-called "insider" perception is still shaped by broader ideological influences. As perceptual dialectology research on linguistic (in)security shows, listeners may align with dominant ideologies even when evaluating their own speech communities. Preston's (1999) perceptual dialectology work illustrates this: Alabamans reported their own speech as "incorrect" but also as the "most pleasant," revealing a dual orientation toward both dominant cultural evaluations and familiar linguistic norms.[1]

3 | Moving Forward: Uniting Sociolinguistic Perception and the Listening Subject

The *listening subject* offers a valuable framework for sociolinguistic perception research in that it foregrounds the need to make explicit the ideological position from which listeners hear and assign meaning (see also Sprenger 2025). Sociolinguistic perception is shaped by both an individual's linguistic background as well as broader hegemonic discourses and institutional ideologies. Recognizing the listening subject therefore means asking not just *who* the listener is, but *what social conditions* make their hearing possible. Edwards (2018) illustrates how this question might be approached through examination of the infrastructure that facilitates communication among DeafBlind people at Gallaudet University. Drawing on Gibson's (1986) notion of "affordances," Edwards observes that "perceptions of what is possible in communication are instances of more general perceptions of what is possible in life. For example, the affordances of a lake's surface are different for a human and a water bug. The way humans and water bugs communicate is also different, and the two facts are not unrelated" (275). Shifting focus toward the *affordances* of the social context offers greater explanatory power by moving beyond observations that a given feature indexes some social meaning to understanding the mechanisms through which that meaning is produced. The key question becomes: Why does this variant mean what it means *to this listener in this moment?* A listening subject framework locates the answer not in the signal itself but in the interpretive position through which that signal is filtered.

We can also consider what such an approach means practically for our research participants—our real-world listening subjects—and how our research designs position them. As Vaughn and Walker (2024) point out, there is often a tension between sociolinguistics' outward-facing message—that all ways of speaking are valid—and the structure of many of our most productive methods, which ask participants to evaluate speakers along various social dimensions. How, then, do we reconcile our goals with our methods? The often decontextualized presentation of speech in experimental sociolinguistic perception studies strikingly resembles the conditions under which the listening subject emerges. Listening participants evaluate voices in the absence of social interaction, often across a spatiotemporal juncture (see Love 2025), and stripped of the possibility of meaning negotiation—a state thatwe might call *experimental eavesdropping*. The listener is positioned as an overhearer, and therefore an outsider, tasked with interpreting speech that is not addressed to them.[2] Yet listeners still make meaning from these isolated fragments—and even converge toward the speech of their imagined interlocutors (Wade 2022; Babel 2012; Nielsen 2011). Listeners' behavior, therefore, can only be interpreted in light of the *listening subject* position in which they have been placed as research participants.

This reframing has important consequences for how we think about public engagement and sociolinguistic outreach. While such efforts have long focused on demonstrating that all language varieties are as systematic and rule-governed as any other, legitimizing stigmatized varieties alone often leaves intact the authority of dominant listening subject positions (see also Clemons 2025). The varieties we teach about are still often presented in contrast to some other imagined "neutral" voice, and the hegemonic position of ideologically privileged voices is exerted through their treatment as unremarkable and unexamined. It is therefore equally important to unsettle this position of auditory privilege by making the accentedness of the imagined "standard" audible for our research participants. For instance, we have begunto explore these dynamics in our lab by asking participants to evaluate the social and linguistic distinctiveness of language varieties perceived by in-group listeners as unmarked or standard—for example, by asking Kansans to identify local speech. Asking participants to locate accent in voices they might otherwise hear as "neutral" can challenge how speakers of these varieties perceive their own "unaccentedness." Further,

interrogating their assumptions about their own speech could prompt recognition that their own way of speaking is no more neutral or correct than any other.

Foregrounding the listening subject invites us to reconsider not only how we interpret perceptual data but how we design our studies and engage with our participants. By making the ideological position of the listener visible, we open space to challenge normative assumptions, unsettle auditory privilege, and invite participants to critically examine their own role in the construction of linguistic meaning.

Endnotes

[1] One of the editors helpfully points out that this dual orientation is akin to the classic status/solidarity divide in work on language attitudes.

[2] One of the editors asks how evaluation would differ if listening participants were positioned as intended addressees. I agree that this is an important consideration in experimental design.

References

Babel, M. 2012. "Evidence for Phonetic and Social Selectivity in Spontaneous Phonetic Imitation." *Journal of Phonetics* 40, no. 1: 177–189. https://doi.org/10.1016/j.wocn.2011.09.001.

Benor, S. B. 2011. "*Mensch*, *Bentsh*, and *Balagan*: Variation in the American Jewish Linguistic Repertoire." *Language & Communication* 31, no. 2: 141–154. https://doi.org/10.1016/j.langcom.2010.08.006.

Bennis, H., and F. Hinskens. 2022. "Coherent Patterns in Nonstandard Inflection in Modern Colloquial Standard Dutch?"" in *The Coherence of Linguistic Communities: Orderly Heterogeneity and Social Meaning*, edited by K. V. Beaman and G. R. Guy, (1st ed.). Routledge. https://doi.org/10.4324/9781003134558.

Campbell-Kibler, K. 2007. "Accent, (ING), and the Social Logic of Listener Perceptions." *American Speech* 82, no. 1: 32–64. https://doi.org/10.1215/00031283-2007-002.

Casasanto, L. S. 2008. "Does Social Information Influence Sentence Processing?" In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 30, edited by B. C. Love, K. McRae, and V. M. Sloutsky. Cognitive Science Society.

Clemons, A. M. 2025. "Excavating the Human in Linguistic Research." *Journal of Linguistic Anthropology*. https://doi.org/10.1111/jola.70025.

D'Onofrio, A. 2015. "Persona-Based Information Shapes Linguistic Perception: Valley Girls and California Vowels." *Journal of Sociolinguistics* 19, no. 2: 241–256. https://doi.org/10.1111/josl.12115.

D'Onofrio, A. 2018. "Controlled and Automatic Perceptions of a Sociolinguistic Marker." *Language Variation and Change* 30, no. 2: 261–285. https://doi.org/10.1017/S095439451800008X.

D'Onofrio, A. 2019. "Complicating Categories: Personae Mediate Racialized Expectations of Non-Native Speech." *Journal of Sociolinguistics* 23, no. 4: 346–366. https://doi.org/10.1111/josl.12368.

Drager, K., and M. J. Kirtley. 2016. "Awareness, Salience, and Stereotypes in Exemplar-Based Models of Speech Production and Perception." In *Awareness and Control in Sociolinguistic Research*, edited by A. Babel, 1–24. Cambridge University Press.

Edwards, T. 2018. "Re-Channeling Language: The Mutual Restructuring of Language and Infrastructure Among DeafBlind People at Gallaudet University." *Journal of Linguistic Anthropology* 28, no. 3: 273–292. https://doi.org/10.1111/jola.12199.

Gregersen, F., and N. Pharao. 2016. "Lects Are Perceptually Invariant, Productively Variable: A Coherent Claim About Danish Lects." *Lingua* 172–173. 26-44. https://doi.org/10.1016/j.lingua.2015.12.001.

Gibson, J. G. 1986. "The Theory of Affordances." In *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, edited by R. Shaw and J. Bransford, 127–143. Wiley.

Hay, J., P. Warren, and K. Drager. 2006. "Factors Influencing Speech Perception in the Context of a Merger-in-Progress." *Journal of Phonetics* 34, no. 4: 458–484.

Inoue, M. 2003. "The Listening Subject of Japanese Modernity and His Auditory Double: Citing, Sighting, and Siting the Modern Japanese Woman." *Cultural Anthropology* 18, no. 2: 156–193. https://doi.org/10.1525/can.2003.18.2.156.

Kockelman, P. 2010. "Enemies, Parasites, and Noise: How to Take up Residence in a System without Becoming a Term in It." *Journal of Linguistic Anthropology* 20, no. 2: 406–421.

Koops, C., E. Gentry, and A. 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* 14, no. 2.

Labov, W., S. Ash, M. Ravindranath, T. Weldon, M. Baranowski, and N. Nagy. 2011. "Properties of the Sociolinguistic Monitor." *Journal of Sociolinguistics* 15, no. 4: 431–463. https://doi.org/10.1111/j.1467-9841.2011.00504.x.

Love, S. 2025. "Languaging As Time Travel." *Journal of Linguistic Anthropology*. https://doi.org/10.1111/jola.70024.

Moore, E., and C. Montgomery. 2018. "The Dialect of the Isles of Scilly: Exploring the Relationship Between Language Production and Language Perception in a Southern Insular Variety." In *Southern English Varieties Then and Now*, edited by L. Wright, 39–73. de Gruyter.

Niedzielski, N. 1999. "The Effect of Social Information on the Perception of Sociolinguistic Variables." *Journal of Language and Social Psychology* 18, no. 1: 62–85. https://doi.org/10.1177/0261927X99018001005.

Nielsen, K. 2011. "Specificity and Abstractness of VOT Imitation." *Journal of Phonetics* 39, no. 2: 132–142. https://doi.org/10.1016/j.wocn.2010.12.007.

Park, B., and M. Rothbart. 1982. "Perception of Out-group Homogeneity and Levels of Social Categorization: Memory for the Subordinate Attributes of in-group and Out-group Members." *Journal of Personality and Social Psychology* 42, no. 6: 1051–1068. https://doi.org/10.1037/0022-3514.42.6.1051.

Preston, D. 1999. "A Language Attitude Approach to the Perception of Regional variety." In *Handbook of Perceptual Dialectology*, edited by D. Preston, 359–373. John Benjamins.

Reyes, A. 2017. "Inventing Postcolonial Elites: Race, Language, Mix, Excess." *Journal of Linguistic Anthropology* 27, no. 2: 210–231. https://doi.org/10.1111/jola.12156.

Sprenger, A. 2025. "Listening at Different Scales: Sociolinguistic Perception and the Listening Subject." *Journal of Linguistic Anthropology*. https://doi.org/10.1111/jola.70023.

Staum-Casasanto, L. 2010. "What Do Listeners Know About Sociolinguistic Variation?" *University of Pennsylvania Working Papers in Linguistics* 15, no. 2: 240–249.

Vaughn, C., and A. Walker. 2024. ""I Always Think That These Are Funny": The Experience of Being a Participant in a Speaker Perception Task." Paper presented at New Ways of Analyzing Variation (NWAV) 52, Miami, FL, November 7–9.

Wade, L. 2022. "Experimental Evidence for Expectation-Driven Linguistic Convergence." *Language* 98, no. 1: 63–97. https://doi.org/10.1353/lan.2021.0086.

Wade, L., T. Hausthor, and J. Holliday. Forthcoming. "What Does Kansas Speech Sound Like? It Depends Who You Ask". *American Speech*.

Wade, L. R., D. Embick, and M. Tamminga. 2023. "Dialect Experience Modulates Cue Reliance in Sociolinguistic Convergence." *Glossa Psycholinguistics* 2, no. 1. https://doi.org/10.5070/G6011187.

Wong, P., and M. Babel. 2017. "Perceptual Identification of Talker Ethnicity in Vancouver English." *Journal of Sociolinguistics* 21, no. 5: 603–628. https://doi.org/10.1111/josl.12264s.