

## Because I'm Not a NAZI? Sociolinguistic Variation and Cross-Linguistic Race Talk

Allegedly, Charles W. Mills once asked Jürgen Habermas: “Why don't you work on race?” Habermas, frowning, replied: “Because I'm not a NAZI?” This exemplifies an interlinguistic misunderstanding that reflects different attitudes toward *race* in the United States and *Rasse* in Germany. Mills's question assumes that a critical theorist has good reason to engage with race. Habermas's response suggests the opposite: only a Nazi would study *Rasse*.

We explore the hypothesis that this and similar exchanges reflect sociolinguistic variation rather than a semantic difference between “race” and “Rasse”. What unsettles many German speakers is not what “Rasse” means but who typically uses it and which perspectives it is heard expressing. In this respect, the word behaves in ways familiar from discussions of slurs. Examining this empirically, we conducted several experiments with German and U.S. American participants.

Our theoretical framework builds on work on slurs and sociolinguistic variation treating lexical choice as evidence of social positioning rather than as a mere vehicle for descriptive content. Nunberg (2016) argues that slurs – and, we suggest, other socially charged terms like *Rasse* – convey information about their users, signalling group affiliation, ideological alignment, and complicity with particular worldviews. Generalising this insight, Nowak (forthcoming) emphasises the role of metadata: socially encoded knowledge about a term's histories of use, typical users, and associated perspectives, which shapes how utterances are interpreted and speakers are evaluated.

This perspective aligns with sociolinguistic accounts that treat linguistic variation as a pragmatic resource for managing social meaning (Burnett 2019). On such views, word choice positions speakers in ways that invite particular evaluations of stance or allegiance, independently of propositional content. Camp (2013) and Bolinger (2015) show how lexical choices can foist perspectives or agendas onto audiences regardless of speakers' explicit intentions, while Davis and McCready (2020) argue that slurs invoke a complex of sociohistorical facts and attitudes even when used by speakers who do not endorse them. Taken together, these accounts support a view on which projective or expressive effects are not exhausted by truth-conditional meaning, but can arise from a term's social provenance and conditions of uptake.

Seen through this lens, the cross-cultural contrast becomes intelligible: for many German speakers, *Rasse* is saturated with Nazi provenance and readily heard as affiliating its users with a racist perspective, whereas in the United States *race* is closely tied to civil-rights discourse, social science, and law, and is often heard as analytically – or even solidaristically – motivated. These divergent patterns of uptake thus reflect not a semantic difference between the terms, but distinct sociolinguistic ecologies that structure how lexical choices are interpreted and evaluated.

In one experiment, German and U.S.-American participants read short vignettes in which a speaker used either *race* or *Rasse* to ask a question that could plausibly be interpreted as racist or as anti-racist/social-scientific. Participants rated the speaker on 7-point scales (racist ... anti-racist; conservative ... liberal). The results revealed a clear intergroup difference: German participants tended to rate the speaker as racist and right-wing, whereas U.S.-American participants did not. Although suggestive, these findings do not by themselves distinguish between a semantic and a metadata-based explanation.

On a semantic explanation, *Rasse* carries racist truth-conditional content (Ludwig 2018; cf. Appiah 1990): its correct application presupposes that humanity is divided into biologically grounded races whose members share inherited traits extending beyond surface morphology and bearing on normatively salient capacities or dispositions. Because even questions carry existential

presuppositions, the use of *Rasse* would thereby commit the speaker to the existence of such races, which may suffice to license the judgement “racist”.

A natural diagnostic for our aim is the P-family (projection) test (Chierchia & McConnell-Ginet 2000), identifying content that survives negation, questioning, or hypothetical embedding. However, both existential presuppositions associated with racist semantic content and socially encoded metadata tied to a term’s provenance can give rise to projection effects. As Tonhauser et al. (2013) emphasise, the P-family test alone cannot determine what kind of not-at-issue content is responsible.

At this stage, our aim is therefore not to classify projective content exhaustively, but to test whether the observed speaker evaluations are better explained by responsibility for lexical choice under conditions of sociolinguistic risk than by truth-conditional semantic commitments alone. Importantly, such effects need not be taken to show that *Rasse* encodes expressive content in the sense of Potts (2007); they can arise from socially salient provenance and asymmetric responsibility for articulation, even where lexical meaning itself remains neutral.

We thus examine embeddings of *Rasse* in direct and indirect speech (Cepollaro et al. 2024). Neither construction commits the reporting speaker to the existence of races: if A reports that B said that C and D belong to different races, or quotes B as saying so, A does not thereby endorse a racial ontology. However, speakers using indirect speech bear responsibility for their own lexical choices. Because indirect reports permit alternative formulations—paraphrase, scare quotes, or explicit distancing—failure to avoid or mark a socially charged term may be taken to signal affiliation with its provenance. By contrast, direct quotation reproduces another’s words rather than selecting one’s own.

We therefore test the interaction between word choice (*Rasse* vs. *Arschloch*) and report type (direct vs. indirect). We predict (i) a significant difference in racism ratings between indirect and direct reports featuring *Rasse*, reflecting the projection of provenance-based metadata under asymmetric responsibility for lexical choice, and (ii) a substantially weaker indirect/direct difference for *Arschloch*—a hybrid pejorative whose negativity is lexically encoded—consistent with its evaluative force being less dependent on social provenance (Milić 2018).

If these predictions are confirmed, the racism ratings observed in our first study can be shown to arise, at least in part, from metadata associated with *Rasse*, independently of any racist semantic content it might encode. Together, the experiments illuminate what was at issue in the Mills–Habermas exchange and help explain why race discourse diverges across linguistic and national contexts: German unease over *Rasse* reflects a tendency to disaffiliate from its perceived primary users rather than a disagreement over biological facts.

## References

- Appiah, Kwame Anthony. 1990. "Racisms." In *Anatomy of Racism*, edited by David Theo Goldberg, 3–17. Minneapolis: University of Minnesota Press.
- Bolinger, Renée Jorgensen. 2015. "The Pragmatics of Slurs." *Noûs* 51 (3): 439–62.
- Burnett, Heather. 2019. "Signalling Games, Sociolinguistic Variation and the Construction of Style." *Linguistics and Philosophy* 42 (5): 419–450.
- Camp, Elisabeth. 2013. "Slurring Perspectives." *Analytic Philosophy* 54 (3): 330–49.
- Cepollaro, Bianca, Simone Sulpizio, Claudia Bianchi, and Isidora Stojanovic. 2024. "Slurs in Quarantine." *Mind & Language* 39 (3): 381–96.
- Chierchia, Gennaro, and Sally McConnell-Ginet. 2000. *Meaning and Grammar: An Introduction to Semantics*. 2nd ed. Cambridge, MA: MIT Press.
- Davis, Christopher, and Eleni McCready. 2020. "The Instability of Slurs." *Grazer Philosophische Studien* 97: 63–85.
- Degen, Judith, and Judith Tonhauser. 2025. "Projection Inferences: On the Relation between Prior Beliefs, At-Issuehood, and Lexical Meaning." *Glossa Psycholinguistics* 4 (1).
- Hom, Christopher. 2008. "The Semantics of Racial Epithets." *The Journal of Philosophy* 105 (8): 416–40.
- Lepore, Ernest, and Una Stojnić. 2024. *On Inflammatory Language*. Oxford: Oxford University Press.
- Ludwig, David. 2018. "How Race Travels: Relating Local and Global Ontologies of Race." *Philosophical Studies* 175: 1–22.
- Milić, Ivan. 2018. "What Counts as an Insult?" *Acta Analytica* 33 (4): 539–52.
- Nowak, Ethan. Forthcoming. "Sociolinguistic Variation, Slurs, and Speech Acts." *Journal of Philosophy*.
- Nunberg, Geoff. 2016. "The Social Life of Slurs." In *New Work on Speech Acts*, edited by Daniel Fogal, Daniel Harris, and Matt Moss. Oxford: Oxford University Press.
- Potts, Christopher. 2005. *The Logic of Conventional Implicature*. Oxford: Oxford University Press.
- Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. "Toward a Taxonomy of Projective Content." *Language* 89 (1): 66–109.

