

Toward Responsible ASR for African American English Speakers: A Scoping Review of Bias and Equity in Speech Technology

Jay L. Cunningham¹, Adinawa Adjagbodjou³, Jeffrey Basoah², Jainaba Jawara⁵, Kowe Kadoma⁴, Aaleyah Lewis²

¹DePaul University — Chicago, IL USA

²University of Washington — Seattle, WA USA

³Carnegie Mellon University — Pittsburgh, PA USA

⁴Cornell Tech University — New York City, NY USA

⁵University of Maryland — College Park, MD USA

jcunni37@depaul.edu, aadjagbo@andrew.cmu.edu, jeffkb28@uw.edu, jjawara@umd.edu, kk696@cornell.edu, alewis9@cs.washington.edu

Abstract

This scoping literature review examines how fairness, bias, and equity are conceptualized and operationalized in Automatic Speech Recognition (ASR) and adjacent speech and language technologies (SLT) for African American English (AAE) speakers and other linguistically diverse communities. Drawing from 44 peer-reviewed publications across Human-Computer Interaction (HCI), Machine Learning/Natural Language Processing (ML/NLP), and Sociolinguistics, we identify four major areas of inquiry: (1) how researchers understand ASR-related harms; (2) inclusive data practices spanning collection, curation, annotation, and model training; (3) methodological and theoretical approaches to linguistic inclusion; and (4) emerging practices and design recommendations for more equitable systems. While technical fairness interventions are growing, our review highlights a critical gap in governance-centered approaches that foreground community agency, linguistic justice, and participatory accountability. We propose a governance-centered ASR life-cycle as an emergent interdisciplinary framework for responsible ASR development and offer implications for researchers, practitioners, and policymakers seeking to address language marginalization in speech AI systems.

Introduction

The rapid growth and widespread application of Automatic Speech Recognition (ASR) systems have opened new possibilities across domains such as transcription services, accessibility tools, and virtual assistants, promising greater access to users. However, these systems continue to demonstrate systemic disparities in performance, particularly for speakers of African American English (AAE), whose linguistic patterns are frequently misrecognized, misinterpreted, or erased by ASR systems predominantly trained on standardized American English (SAE) datasets (Koencke et al. 2020; Mengesha et al. 2021). These failures are not merely technical errors; they reflect and reinforce broader societal inequities around whose voices are recognized, valued, and legitimized by AI systems (Wenzel and Kaufman 2023;

Brewer, Harrington, and Heldreth 2023; Hanna et al. 2020; Smith et al. 2020).

These harms also reflect long-standing language ideologies that shape which forms of speech are deemed legitimate in society. Language variation, has long been a focus in sociolinguistics, which has shown how dominant attitudes toward non-standard dialects and vernaculars uphold social hierarchies and racialized perceptions. In particular, sociolinguistic scholarship has investigated how language upholds social inequalities and perpetuates bias specifically for AAE speakers (Labov 1963; Rosa and Flores 2017; Morgan 2002; Green 2002a) despite decades of significant sociolinguistic efforts to discredit discriminatory narratives and work to legitimize it as a systematic, rule-governed variety (Labov 1963; Poplack 1980). These views, rooted in deficit-based ideologies (Rosa and Flores 2017; Rickford 2016), persist in language technologies, where users who do not conform to dominant language practices (Blodgett et al. 2020) often face negative consequences.

For AAE speakers, ASR system breakdowns reproduce historical patterns of exclusion, erasure, and discrimination embedded in broader algorithmic and data-driven harms (Wenzel et al. 2023; Bird 2024). These harms manifest as usability failures, trust erosion, and miscommunication, but also extend into moderation, surveillance, and policing contexts, where mis-recognition can have direct material and safety consequences (Mengesha et al. 2021; Cunningham et al. 2024). These systemic harms highlight the urgent need to re-evaluate the ASR development pipeline itself, particularly the universal design assumptions that erase linguistic and cultural specificity.

In this scoping review (Tricco et al. 2018), we focus explicitly on ASR as a speech AI technology, recognizing that while our findings also inform broader Speech and Language Technologies (SLTs) and Natural Language Processing (NLP) systems, the primary site of harm, exclusion, and intervention discussed herein remains ASR systems deployed in U.S. English-language contexts serving Black AAE speakers. Addressing these disparities requires an interdisciplinary synthesis of research practices, data methods, and governance frameworks across HCI, ML/NLP, and so-

ciolinguistics, as efforts in these domains have often been fragmented and siloed. While AAE is the focal case for this review, given its sociolinguistic complexity, historical marginalization, and documented ASR disparities, the findings and governance-centered framework are intended to serve as a transferable model for advancing inclusive ASR practices across other underrepresented dialects and languages. This framing ensures the work contributes both to the specific remediation of harms in AAE contexts and to the broader theoretical and practical discourse on linguistic equity in speech AI.

While prior literature reviews, such as Ngueajio and Washington (2022) survey on ASR system biases and mitigation techniques, have synthesized technical approaches to bias identification and correction, these works primarily center model-level interventions and error mitigation strategies. In contrast, this review adopts a broader socio-technical and governance-centered lens, treating ASR fairness not solely as a matter of model optimization but as a systemic issue rooted in data practices, linguistic marginalization, and exclusionary design norms. Our work extends beyond system performance to interrogate whose linguistic practices are privileged, how epistemic authority is distributed in ASR development, and what participatory and community-led models of design and evaluation are emerging or still lacking. By foregrounding linguistic justice, cultural legitimacy, and the governance of ASR systems, this review addresses critical gaps in existing scholarship and contributes to ongoing conversations on decolonial computing and algorithmic harm reduction.

This review seeks to bridge these gaps by synthesizing existing interdisciplinary scholarship, highlighting tensions and synergies across fields, and offering pathways toward more accountable, participatory, and community-centered ASR development.

Research Questions This study is guided by the following research questions (RQs):

- RQ1: Across Human-Computer Interaction (HCI), ML/NLP, and Linguistics, how are researchers discussing and addressing fairness, bias, and equity in ASR for Black speakers of AAE?
- RQ2: What are the applied data practices (curation, collection, validation, annotation, deployment) being used or proposed in ASR systems that aim to be inclusive of AAE speakers and Black users?
- RQ3: What broader conceptual design approaches, methods, and theories have been employed to advance inclusive ASR systems for Black AAE speakers?
- RQ4: What are practical recommendations for designing, deploying, or governing more AAE-inclusive ASR systems?

These research questions seek to document the various intervention points where researchers, practitioners, and other stakeholders engage with ASR development, and to assess the tangible outcomes these actions have yielded. Beyond technical practices alone, we take an intentional approach to documenting discourse, practices, and recommendations,

with the aim of capturing the expansive strategies that future researchers can build upon to contribute to more accountable, linguistically-inclusive, and community-centered ASR technologies. We focus on AAE as an illustrative case of equity challenges in speech technology, addressing urgent social justice concerns while generating transferable insights for inclusive ASR design across other minoritized dialects and languages.

With this work, we contribute:

1. An interdisciplinary synthesis of the current responsible ASR landscape, critically analyzing how fairness, bias, and equity are addressed across HCI, ML/NLP, and linguistics, with a specific focus on the lived experiences and speech practices of Black AAE speakers.
2. A mapping of data practices, design approaches, and evaluation methods currently used or proposed to mitigate ASR disparities, highlighting tensions, synergies, and disciplinary blind spots.
3. A governance-centered reflection on the limitations of existing approaches, offering pathways toward more participatory, accountable, and community-centered ASR development that explicitly challenges universal design assumptions and centers linguistic justice.

This work advances the AIES agenda by foregrounding the systemic and socio-technical harms embedded in ASR systems, emphasizing that addressing these issues requires not only technical interventions but also structural changes in who governs, designs, and evaluates ASR technologies, and how those processes can be re-imagined to center the voices and agency of Black AAE speakers.

**An extended version of this paper, which includes supplementary materials, technical appendices, and expanded methodological details, is available on arXiv¹ under a Creative Commons Attribution (CC BY) license.*

Background

To situate our scoping review, this section examines how language variation, particularly AAE, intersects with technical infrastructures and social power in ASR. We draw from sociolinguistics, HCI, ML/NLP, and speech engineering to outline the historical and contemporary barriers to inclusive ASR development and position the need for governance-centered intervention.

Language Variation and Socio-technical Hierarchies

Language is not merely a technical input but a socially encoded practice shaped by ideology, identity, and inequality. Sociolinguistic studies have long demonstrated that linguistic variation across dialects, regions, and communities is both systematic and meaningful (Labov 1963; Green 2002a). Yet, dominant ideologies frame non-standard varieties, especially AAE, as deficient, contributing to racialized hierarchies in education, employment, and increasingly,

¹navigate to <https://jaylcunningham.com/work> for the arXiv web-link.

AI systems (Rosa and Flores 2017; Rickford 2016). These views persist despite decades of research affirming AAE’s grammatical structure and internal consistency (Labov 1972; Poplack 1980; Lanehart 2015).

This foundational work is critical for contextualizing how bias in ASR systems is not an isolated technical artifact, but the continuation of social and linguistic exclusion via algorithmic infrastructures.

African American English and Inclusive ASR

African American English (AAE), commonly referred to as African American Vernacular English (AAVE) or African American Language (AAL), is a dynamic, rule-governed language variety with deep cultural and historical significance. While early sociolinguistic studies focused narrowly on the speech of urban, working-class Black men, contemporary research demonstrates rich intra-group variation across gender, geography, and age (King 2020; Farrington 2019; Forrest and Wolfram 2019). The fluidity of AAE, shaped by regional, generational, and slang variation, poses challenges for ASR, as standardized training data often miss its shifting phonological, lexical, and stylistic patterns, reducing accuracy and risking intra-community inequities without adaptive, continually updated models (Koenecke et al. 2020; Blodgett, Green, and O’Connor 2016).

Despite this linguistic legitimacy, AAE continues to be stigmatized in language technologies. Studies across HCI, NLP, and ASR consistently show that AAE speakers face higher word error rates, reduced system reliability, and misrecognition of culturally specific language (Koenecke et al. 2020; Mengesha et al. 2021; Harrington et al. 2022). These failures are not neutral but reflect socio-technical design decisions that privilege SAE as the norm.

Attempts to improve system performance for AAE have included dialect-specific datasets (Blodgett, Green, and O’Connor 2016), syntactic disambiguation tools (Santiago et al. 2022), race-primed annotation protocols (Sap et al. 2019), and participatory design methods (Kim et al. 2022). While these interventions are valuable, they often emerge within siloed disciplines, limiting the field’s ability to enact structural change.

Bias and Disciplinary Gaps in ASR Research

Technical disciplines such as ML and speech engineering have made progress on bias mitigation, data curation, and fairness audits (Hutiri and Ding 2022; Xu et al. 2021), but often lack engagement with the sociocultural dynamics that shape language marginalization. Conversely, HCI and sociolinguistic studies center user perspectives, cultural context, and community agency, but are underrepresented in mainstream ASR pipelines.

Recent studies have surfaced racial bias and underrepresentation in ASR systems (Ngueajio and Washington 2022; Blodgett et al. 2020), yet few bridge insights across disciplines. This fragmentation obscures shared goals and undermines opportunities for holistic, equity-centered innovation. Our review addresses this gap by synthesizing interdisciplinary approaches to bias, equity, and fairness in ASR systems for AAE speakers.

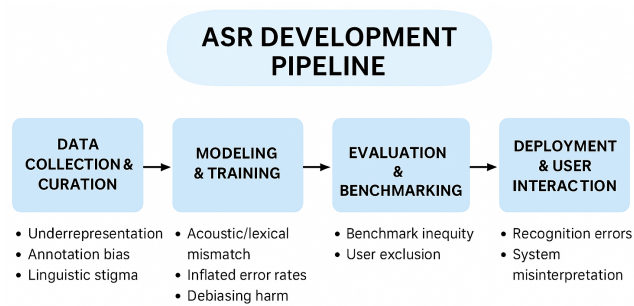


Figure 1: ASR Modeling Development Pipeline and Sites of Harm

ASR Development Pipeline and Sites of Harm

ASR systems function through a multi-step pipeline, data collection, annotation, modeling, evaluation, and deployment, each of which encodes socio-technical values (see Figure 1). At every stage, AAE speakers are vulnerable to exclusion, mis-recognition, or epistemic erasure.

- **Data Collection & Curation:** Corpora often underrepresent AAE or homogenize its variation, leading to under-performing models and biased assumptions (Dorn 2019; Markl and McNulty 2022).
- **Annotation:** Annotators unfamiliar with AAE may mislabel nonstandard forms or reinforce toxicity stereotypes (Sap et al. 2019; Davidson, Bhattacharya, and Weber 2019).
- **Model Training:** Standard acoustic and language models penalize features common in AAE, resulting in elevated error rates and interpretive failures (Koenecke et al. 2020; Okpala et al. 2022).
- **Evaluation:** Benchmarks calibrated to SAE ignore dialectal differences, erasing AAE-specific challenges from performance metrics (Wenzel et al. 2023).
- **Deployment:** Failures in high-stakes domains like healthcare or employment magnify harm and diminish user trust (Crawford 2021).

By mapping these harms across the ASR life-cycle, our study reframes ASR not just as a technical pipeline, but as a governance challenge, demanding intervention at multiple sites of design and accountability. ASR’s failures for AAE speakers illustrate a broader crisis in language AI: the reproduction of racial and linguistic hierarchies through automation. By articulating the socio-technical dimensions of harm and highlighting underexplored disciplinary gaps, this review invites a more justice-centered vision for speech technologies, one that moves beyond representational diversity toward equity-centered governance and design.

Scoping Review Methodology

This review adopts the PRISMA Scoping Review (PRISMA-ScR) methodology as outlined by (Tricco et al. 2018), which is specifically suited for comprehensively mapping emerging and heterogeneous research fields

such as ASR. This review is based solely on secondary analysis of published literature and does not include new empirical testing of ASR systems. Given the interdisciplinary nature of ASR scholarship, this review employs an inductive scoping approach to capture the diversity of epistemological positions, data practices, and design approaches present across disciplines, including HCI, ML/NLP, and sociolinguistics, while remaining open to underrepresented perspectives, such as those centering Black AAE speakers. Scoping reviews are particularly appropriate for fields where the research landscape is complex, fragmented, and under-reviewed, as they provide a structured yet exploratory assessment of the extent, range, and nature of research activity. By applying this method, we aim to map and synthesize the interdisciplinary discourse on ASR fairness and equity, identify knowledge gaps, and inform future research and governance directions.

Inclusion and Exclusion Criteria

This inclusion and exclusion criteria framework guide our literature review process.

Inclusion Criteria Studies are included if they meet any of the following conditions: 1) Focus specifically on AAE or similar American English linguistic variations, 2) Have a strong thematic focus on bias identification, inclusive ASR/NLP design methods, and practical recommendations for inclusive design, aimed at improving the fairness and inclusion of ASR technologies.

Exclusion Criteria Studies are excluded if they: 1) Do not focus on AAE or similar ethnic minority linguistic variations of American English, 2) Are primarily focused on other languages or general ASR technologies without a clear emphasis on inclusivity, equity, or responsible design practices.

Information Sources and Search Strategy To guide our interdisciplinary literature search, we first outlined six thematic clusters reflecting key conceptual areas of fairness and equity in ASR systems: (1) Responsible Methods & Design Approaches, (2) Social & Algorithmic Justice Challenges, (3) Speech & Language Technologies (ASR Focus), (4) Language Varieties & Dialects (U.S. Context), (5) Racial & Ethnic Communities, and (6) Language Group Associations. These clusters were developed to ensure conceptual breadth and to capture both technical and sociolinguistic framings relevant to ASR fairness.

Then, we developed an initial seed set of keywords informed by prior work in AI fairness, sociolinguistics, and speech technology (e.g., “algorithmic bias,” “dialect recognition,” “African American English,” “ASR fairness,” “voice equity,” “linguistic inclusion”). We then broadened our search terminology using ResearchRabbit’s citation network mapping to broaden the conceptual landscape and avoid disciplinary blind spots. This tool generated a network of 1,872 semantically adjacent and related terms, expanding coverage beyond initial expert-defined keywords and surfacing domain-specific variants (e.g., “voice bias mitigation,” “vernacular speech tech,” “speech dialect equity”). This expan-

sion process ensured coverage across ML/NLP, HCI, and sociolinguistics while capturing emergent terminology².

Data Charting Process

Extraction of Evidence Initially, 72 articles were identified and cataloged using Airtable. After applying the inclusion and exclusion criteria, 44 papers were selected for further analysis. These papers specifically focus on inclusive design principles within ASR technologies. The final sample of 44 papers reflects the narrow scope of research explicitly engaging with fairness, bias, or equity in ASR for African American English or comparable non-standard American English varieties. While the number may appear small, it underscores the limited scholarly attention to ASR equity for marginalized speech communities and the need for targeted research in this area.

Charting of Evidence³ The selected 44 papers were processed using Atlas.ti, a qualitative data analysis software, to facilitate coding. A preliminary codebook was created based on the research questions, employing open coding for emergent themes related to bias detection, inclusive design strategies, and diverse linguistic representation (see Figure 2). This systematic approach allowed for a structured analysis of the collected data, ensuring that all relevant themes were thoroughly explored and documented.

Disciplinary Mapping To assess the interdisciplinary landscape of fairness and equity in ASR systems, we classified the 44 reviewed papers by their primary disciplinary orientation based on publication venue, author expertise, and topical focus. The majority of papers (n=31) were situated within Machine Learning and Natural Language Processing (ML/NLP), reflecting the technical emphasis on bias mitigation, data practices, and model evaluation in speech and language technologies. Human-Computer Interaction (HCI) contributed 11 papers, centering on user experience, participatory design, and socio-technical impacts of ASR systems, particularly for marginalized communities. Linguistics was represented in 2 papers, focusing on the sociolinguistic characteristics of AAE and their implications for ASR performance. This distribution underscores the predominance of technical approaches in the current literature, while highlighting emerging intersections with HCI and linguistics that offer critical perspectives on ASR equity.

Data Items The following details were charted for each article where applicable: Author(s); Year of publication and venue (e.g. CHI, 2023); Paper Title; Origin/country of origin (where the study was published or conducted); Aims/purpose and research questions; Study population or applicability to ethnic and linguistic minority American English speaking groups; The type of study (conceptual or empirical); Methodology

²(See the Appendices (B) in extended paper version for expanded details of representative terms for each cluster.)

³(See the Appendices (B) in extended paper version for the full table of Reviewed Literature organized by Citation, Discipline, and Information Source.)

Data Analysis Team The data was coded by a team of five reviewers, all of whom are experienced researchers with diverse backgrounds in HCI, Computer Science (AI/ML), NLP, and Linguistics (Computational and Social). This diversity in expertise enhances the reliability and breadth of the analysis.

We employed an inductive coding approach using Atlas.ti, guided but not constrained by the research questions. A collaboratively developed codebook emerged through iterative cycles of reading and open coding across the full texts of included papers. Two authors independently coded each paper in each round of reviews, and inter-coder agreement was calculated using *pairwise* Cohen’s kappa (Cohen 1960), averaged across all coder pairs, at each iteration. With six coders annotating the full set ($N = 44$ items), this involved computing κ for each of the $\binom{6}{2} = 15$ coder pairs and taking the mean. Inter-coder reliability for the initial round of full-text coding was $\bar{\kappa} = 0.82$, indicating strong agreement. Disagreements were resolved through group discussion, and the final codebook was applied with 100% consensus across all included papers. The resulting thematic structure (see Section reflects both the empirical diversity of reviewed works and the normative framing of responsible ASR design embedded in our Research Questions.

Synthesizing of Results The synthesis process aggregated the findings from these studies, identifying common patterns and themes as well as gaps in the research. This analysis helped in understanding how responsible design practices are conceptualized and implemented in ASR systems for speakers of AAE, identifying key areas for future research.

Summary of Data The table below shows the number of documents retrieved from various information sources (see Table 1)

Elements	Details (count)
Search Keywords	”Algorithmic Fairness”
Limit To	Scholarly articles (peer-reviewed or otherwise)
ACLWeb	18
ACM Digital Library	19
IEEE Xplore	4
arXiv	1
PNAS (National Academies)	1
Frontiers	1
ProQuest	1

Table 1: Summary Search Protocol

Scoping Review Summary and Findings

Following PRISMA-ScR (Tricco et al. 2018) for the reporting of scoping review results, we first present the process of selecting sources of evidence. We then describe the characteristics of these sources of evidence in terms of publication venue and year. And finally, we present the synthesis of the results answering each research question.

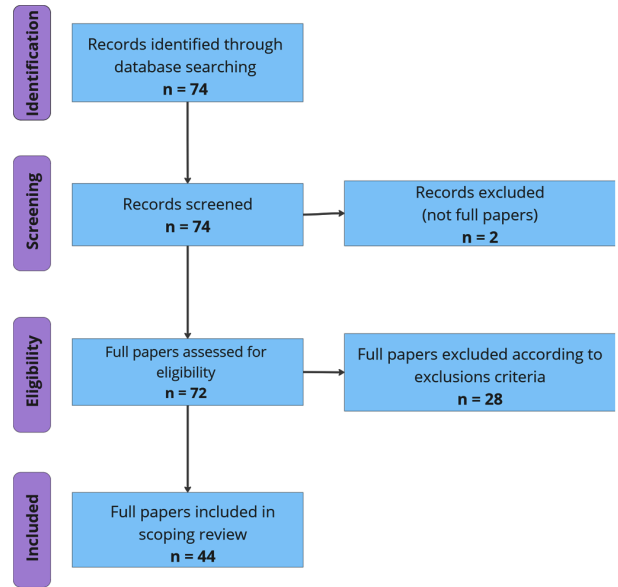


Figure 2: Scoping Review Flow Diagram following the PRISMA method for scoping reviews.

Framing Fairness and Harm in ASR: Disciplinary Perspectives from HCI, ML/NLP, and Linguistics

This section addresses research question 1. We organize findings by disciplinary emphasis, highlighting how each field frames fairness, bias, and equity, and synthesizing their distinct contributions and blind spots.

ML/NLP: Quantifying Disparities and Model Fairness

In ML/NLP, researchers primarily approach ASR fairness through performance evaluation and bias mitigation metrics. Koenecke et al. (2020) document racial disparities in ASR word error rates using commercial systems, revealing that systems perform disproportionately poorly for Black speakers. Sari, Hasegawa-Johnson, and Yoo (2021) propose counterfactual fairness interventions in ASR models, but do not explicitly account for dialectal variance in AAE. Xu et al. (2021) and Ziems et al. (2022) demonstrate that dialect disparity extends beyond ASR into language understanding tasks, showing how minority dialects are marginalized in pretraining and classification phases. Manzini et al. (2019) and Davidson, Bhattacharya, and Weber (2019) uncover bias in hate speech datasets and embeddings, often embedding stereotypes that affect downstream fairness.

Other works explore fairness in speech data augmentation or annotation. Johnson et al. (2022) propose LPC-based augmentation to support underrepresented children’s dialects, while Okpala et al. (2022) use adversarial training to debias hate speech detection. However, despite advances in model-based solutions, few works in this domain consider participatory definitions of fairness or sociolinguistic legitimacy of dialectal inputs. Fairness is often operationalized via demographic parity or performance parity, without scrutiny of epistemic assumptions or linguistic ideologies embedded in models.

HCI: Trust, Experience, and Interaction Harms HCI studies frame ASR fairness through trust, user experience, and emotional impact. Wenzel et al. (2023) and Wenzel and Kaufman (2023) find that ASR, mis-recognition can function as a microaggression, especially across race-aligned interactions. Harrington et al. (2022) examine older Black adults' health information-seeking behaviors, finding that ASR failures prompt code-switching, self-silencing, and distrust. These findings emphasize the affective burden and social exclusion that arise from ASR unreliability.

Accessible interface work such as Lister et al. (2020) and Kim et al. (2022) proposes inclusive design strategies, but does not often evaluate performance differentials along race or dialect lines. Cunningham (2023) calls for community-collaborative approaches to ASR development with AAE speakers to address these lived harms. However, most HCI research does not quantify disparities or directly influence model-level design decisions, indicating a methodological disconnect between user experience research and ML fairness interventions.

Linguistics and Sociolinguistics: Structural Marginalization and Language Ideology Linguists and sociolinguists emphasize how ASR systems encode structural linguistic inequities. Blodgett, Green, and O'Connor (2016) and Rickford (2016) argue that AAE is systematically devalued in both NLP and speech datasets. Bird (2024) critiques the colonial legacies of linguistic extraction, warning that even fairness-oriented inclusion can reinforce linguistic imperialism. Bourdieu's (1991) theories of linguistic capital are used by Weston (2021) and Rosa and Flores (2017) to show how ASR systems reinforce dominant language ideologies.

Recent work by Markl and McNulty (2022) and Markl (2022) highlights how ASR practitioners act as language managers, often enforcing standard language norms while overlooking dialect-specific error patterns. Wassink, Gansen, and Bartholomew (2022) examine acoustic alignment failures with ethnicized dialects in ASR, revealing mismatches in phonetic modeling. Santiago et al. (2022) propose improved modeling of morpho-syntactic features in AAE such as habitual "be," arguing for dialect-sensitive annotation and modeling.

Cross-Cutting Synthesis: Toward Intersectional and Participatory Fairness Across disciplines, researchers converge on the finding that ASR systems disproportionately fail minoritized speakers, particularly those using AAE, through high error rates, coerced linguistic assimilation, and erasure of dialectal legitimacy. Yet, few studies bridge technical, interaction, and ideological dimensions. ML/NLP largely quantifies disparities; HCI explores experience and harm; linguistics interrogates ideological assumptions. Interdisciplinary synthesis remains limited.

Notably, participatory approaches are largely absent. While some HCI and sociolinguistic studies advocate for community-centered design (Cunningham 2023; Harrington et al. 2022), no papers in ML/NLP implement participatory fairness definitions or community-led audits. This gap suggests a critical opportunity for future research: integrating intersectional, participatory, and governance-aware prac-

tices across the full ASR development life-cycle.

Data Practices for Inclusive ASR Design

This section addresses research question 2. We analyze reviewed studies through four major stages of the data life-cycle: collection, curation, annotation, and model training.

Data Collection and Community Participation Inclusive data collection practices for AAE vary across domains, from scraping social media corpora to crowdsourcing speech. In NLP, researchers have collected text from Twitter and other user-generated platforms to compile AAE datasets (Groenwold et al. 2020; Blodgett, Green, and O'Connor 2016; Jørgensen, Hovy, and Søggaard 2015). These datasets often provide linguistic diversity for fine-tuning models but risk entrenching biases if demographic provenance is unclear.

Speech-based collection efforts involve more participatory methods. Markl and McNulty (2022) and Nkemelu et al. (2023) describe community-driven and paid-volunteer approaches to low-resource speech data collection. These methods not only increase linguistic representation but foster community agency in shaping how data is interpreted, especially for hate speech and sociocultural content. However, concerns remain around consent, data commodification, and ownership in crowdsourced platforms.

Curation and Dataset Integration Responsible data curation practices appear across HCI and NLP. Researchers curate corpora that document AAE variation, such as the Corpus of Regional African American Language (CORAAL) (Thomason 2021) and IViE (Grabe, Nolan, and Farrar 1998), which have been used in downstream ASR systems (Farrington 2019). These curated datasets feature transparent regional tagging and transcription practices, facilitating dialect-aware modeling.

Some works combine heterogeneous datasets to augment underrepresented forms of AAE. For example, Santiago et al. (2022) and Johnson et al. (2022) integrate multiple sources for robust modeling. Others apply pre-processing methods like punctuation stripping (Dorn 2019) or more complex adversarial auto-encoding for debiasing (Sari, Hasegawa-Johnson, and Yoo 2021). Manzini et al. (2019) propose a post-curation debiasing step that removes analogical bias from word embeddings before training. These methods reflect growing interest in integrating fairness into data preparation, though sociolinguistic validation is often limited.

Annotation and Fairness Sensitivity Annotation is a key site where bias and subjectivity enter ASR pipelines. Bender and Friedman (2018) and Poletto et al. (2021) argue for documenting annotator demographics and training to mitigate interpretive bias. Several studies report using AAE speakers as annotators to assess or validate interpretations of AAE texts and speech. Groenwold et al. (2020) and Sap et al. (2019) show that race-primed annotators are less likely to label AAE tweets as offensive, suggesting that cultural and racial alignment improves interpretive fairness.

Participatory annotation remains rare but promising. Community-informed schema or iterative consensus models could support more epistemically just annotation practices, especially when documenting sociolinguistic nuance or culturally grounded intent.

Training and Dialect-Specific Modeling Inclusive ASR development also requires training practices attentive to dialectal features. Okpala et al. (2022) retrain BERT on over one million AAE tweets to produce AAEBERT, a hate speech detection model sensitive to AAE syntax and usage. Similarly, Santiago et al. (2022) propose a framework to disambiguate habitual vs. non-habitual uses of “be” in AAE using rule-based filters, while Ziems et al. (2022) construct orthographic variants through morphosyntactic transformation.

Such techniques show promise for recognizing linguistic specificity, but raise questions about over-generalization and ethical representation. Models trained with dialectal precision may still misinterpret intent or meaning without contextual grounding. Few of the reviewed studies integrate human feedback loops or community validation post-training.

Overall, across the data life-cycle, efforts to include AAE speakers in ASR development remain fragmented. While representational improvements are emerging in collection and modeling, participatory governance, community consent, and epistemic alignment remain underdeveloped. This suggests a need for socio-technically integrated data practices which embed linguistic, cultural, and ethical considerations throughout the pipeline.

Approaches to Studying Inclusive ASR and NLP

This section addresses research question 3. We categorize the reviewed literature into three overarching methodological paradigms: quantitative, qualitative, and theoretical.

Quantitative Approaches: Measuring Disparity and Model Intervention Quantitative approaches dominate the reviewed literature, often aiming to measure or mitigate disparities through statistical or computational means. Core methods include word error rate (WER) analysis, comparative classifier evaluation, sentiment scoring, and debiasing.

Several studies quantify model-level harms: Koenecke et al. (2020), Dorn (2019), and Johnson et al. (2022) use WER to expose transcription disparities between AAE and SAE, while Markl and McNulty (2022) employs regression analysis to model algorithmic bias in British English ASR. Wenzel et al. (2023) and Davidson, Bhattacharya, and Weber (2019) use statistical analysis to assess psychological effects and classification bias. Groenwold et al. (2020) conducts BLEU/ROUGE comparisons and sentiment analysis on GPT-2 outputs for AAE and SAE.

Debiasing and detoxification methods offer technical mitigation pathways. Xu et al. (2021) explores four strategies to guide generative models away from toxic outputs, while Manzini et al. (2019) proposes a multiclass embedding debiasing algorithm. Framework-guided quantitative work includes Joshi et al. (2020), who builds a language resource disparity taxonomy, and Hutiri and Ding (2022), who pro-

poses a bias quantification framework for speaker verification challenges.

While these approaches offer reproducible measures of disparity, they often overlook community-defined fairness goals and fail to interrogate how inclusion is conceptually defined.

Qualitative and Participatory Approaches: Centering Lived Experience and socio-technical Context Qualitative studies investigate inclusive ASR through user-centered and community-engaged lenses. Common methods include interviews (Kim et al. 2022), surveys and diary studies (Mengesha et al. 2021; Wenzel et al. 2023), and observation (Iacobelli and Cassell 2007). These approaches capture affective, psychological, and sociocultural responses to ASR failures, particularly among Black and AAE-speaking users.

Participatory design methods, spanning co-design (Kim et al. 2022), community-based design (Racadio, Rose, and Kolko 2014), and cooperative inquiry (Björgvinsson, Ehn, and Hillgren 2010), emerge across several HCI-oriented studies. Ziems et al. (2022) and Nkemelu et al. (2023) apply expert-informed validation protocols and context-sensitive annotation, respectively, to guide benchmark or model development. Others use participatory annotation strategies to better preserve dialectal meaning or assess offensive content interpretation (Groenwold et al. 2020; Sap et al. 2019; Jørgensen, Hovy, and Søggaard 2015).

These methods not only enhance design inclusion but function as socio-technical governance mechanisms, redefining who participates in NLP knowledge production and how authority is distributed. Yet, most remain confined to design phases and are rarely integrated into upstream model development or downstream evaluation.

Theoretical Frameworks: Critical Lenses on Power, Language, and Identity A smaller subset of studies ground their analysis in theory-driven inquiry. Markl and McNulty (2022) applies intersectionality to trace compounding harms across racial, linguistic, and geographic dimensions. Sutton et al. (2019) introduce sociophonetic-inspired design strategies that highlight context-awareness and voice individualization. Bird (2024) adopts a decolonial approach to contest extractive language technologies and advocate for Indigenous sovereignty over linguistic data.

Frameworks such as microaggressions and harm-based equity (Wenzel et al. 2023), and Critical Race Theory (Crenshaw 1991), guide ethical critique and propose reparative justice principles. These works surface epistemic injustice in NLP and call for more reflexive and politicized design practices. However, such frameworks are rarely operationalized into data, modeling, or evaluation standards.

In sum, while diverse methodological traditions are represented in the literature, most approaches remain siloed. Quantitative studies dominate, but often neglect cultural nuance or participatory principles. Qualitative and theoretical work illuminate structural harms and lived experience but lack integration into dominant AI pipelines. Future research must advance mixed-method and theory-informed approaches that foreground the agency of marginalized speakers in defining and achieving fairness.

Practical Recommendations for Inclusive ASR/NLP Design

To explore research question 4, this section synthesizes practical strategies for ASR development.

Community Engagement and Problem Framing Designing inclusive systems begins with centering communities in defining problems, goals, and design values. Research in HCI and sociolinguistics underscores the need to co-develop technology with the linguistic communities it is meant to serve (Kim et al. 2022; Cunningham 2023; Bird 2024). These studies engage AAE speakers through participatory design and interviews, surfacing culturally relevant use cases, frustrations with existing systems, and desired features. Scholars advocate reframing speech tech development as civic infrastructure (Markl and McNulty 2022), where public input, trust, and ethical alignment guide innovation from the outset (Mengesha et al. 2021; Brewer, Harrington, and Heldreth 2023).

Ethical and Inclusive Data Collection Inclusive ASR/NLP development depends on intentional, consent-based data collection that reflects community norms and linguistic realities. Researchers emphasize collecting speech data from diverse AAE speakers using both social media corpora (Groenwold et al. 2020; Blodgett, Green, and O'Connor 2016) and community-centered efforts like Project Elevating Black Voices, a partnership where data sovereignty remains with historically Black institutions (HBCUs) (Heldreth 2023). In low-resource contexts, community-led data practices, such as those by Māori speakers or hate speech annotators in Sub-Saharan Africa, provide models for ethical stewardship (Nkemelu et al. 2023; Brown et al. 2024).

Responsible Data Curation and Annotation Annotation must be approached as a socio-technical process shaped by language ideologies and annotator positionality (Bender and Friedman 2018; Papakyriakopoulos and Xiang 2023). AAE-inclusive projects often recruit code-switching annotators or apply dialect/race priming to mitigate labeling bias (Sap et al. 2019; Groenwold et al. 2020). Situated annotation practices, matching annotators to speaker demographics, help preserve meaning and reduce misidentification (Papakyriakopoulos and Xiang 2023). These approaches demand transparency around annotator training and a commitment to dialectal accuracy.

Dataset Design Dynamics and Representation Inclusive datasets go beyond quantity to account for representational ethics, balance, and diversity (Jo and Gebru 2020). Scholars propose curating datasets by dialect, region, and socio-cultural attributes (Mengesha et al. 2021; Markl and McNulty 2022), and balancing linguistic data across socio-demographic strata (e.g., age, gender, socioeconomic status) (Wassink, Gansen, and Bartholomew 2022). Assets-based approaches shift the design lens from deficit to strength, affirming the linguistic value of marginalized varieties (Wenzel et al. 2023; Harrington et al. 2022).

Inclusive Model Training and Evaluation Model development must accommodate linguistic variation while pro-

tecting communities from harm. Techniques such as dialect-specific pretraining (e.g., AAEBERT) (Okpala et al. 2022), rule-based transformations (Ziems et al. 2022), and adversarial filtering (Sari, Hasegawa-Johnson, and Yoo 2021) offer concrete practices for mitigating dialectal bias. Scholars argue that WER-based metrics must be augmented with qualitative evaluations, including microaggression impacts and user trust (Koenecke et al. 2020; Wenzel et al. 2023). Others propose dynamic evaluation thresholds by speaker group to minimize predictive harms (Hutiri and Ding 2022).

Deployment, Customization, and Accountability Deployment strategies should promote user autonomy and system responsiveness. Context-aware and individualized designs allow users to select preferred voices and correct misrecognitions (Sutton et al. 2019; Mengesha et al. 2021). ASR systems can be designed to issue affirmations or apologies following recognition failures, redistributing accountability to the system rather than the user (Wenzel et al. 2023). Researchers propose community-based review panels or impact assessments to ensure ethical use in high-stakes contexts (e.g., health, education, policing) (Blodgett et al. 2022).

Reflexive and Interdisciplinary Evaluation Finally, inclusive ASR/NLP development must incorporate mechanisms for continuous reflection and adaptation. This includes positionality statements (Allen 1996), diversity audits, and community accountability protocols (Joshi et al. 2020; Sutton et al. 2019). Sociolinguistics and raciolinguistics provide critical tools for interpreting model behavior and social impact (Rickford 2016; Wenzel and Kaufman 2023). Future work should build interdisciplinary teams and embed equity goals throughout project governance, from funding to peer review.

In sum, the inclusive ASR/NLP life-cycle framework outlined here moves beyond piecemeal fixes to articulate a governance-centered, community-rooted agenda for language technology. These interventions are not just technical, they are socio-technical, epistemic, and political. By reimagining development through these lenses, this review offers scholars and practitioners a scaffold for building systems that affirm, rather than erase, the linguistic richness of marginalized communities.

Discussion

This scoping review analyzed 44 interdisciplinary publications across ML/NLP, HCI, and sociolinguistics to assess how fairness, bias, and equity are addressed in ASR systems for Black speakers of AAE. Rather than applying a pre-existing analytic lens, we adopted an inductive synthesis approach, allowing critical insights to emerge through comparative literature analysis. One striking insight is the absence of governance-oriented interventions across most of the reviewed literature. While technical fairness auditing and inclusive data practices are present, few studies interrogate how power, accountability, and community agency shape ASR system outcomes. This finding signals a significant gap and opportunity to situate fairness socio-technical commitment to be governed.

A Framework for Governance-Centered ASR

The findings from our scoping review directly motivate the design of the governance-centered ASR life-cycle. Despite progress in bias detection and data diversification, current fairness interventions in ASR overwhelmingly focus on narrow technical fixes, model tuning, dataset re-balancing, and retrospective evaluations of word error rates (Sap et al. 2019; Koenecke et al. 2020). These interventions are important but insufficient. They often abstract fairness into quantifiable metrics without attending to the underlying political, social, and epistemic power structures that shape how ASR systems are developed, deployed, and governed (Blodgett et al. 2020; Madaio et al. 2020). In response, we propose a Governance-Centered ASR life-cycle as a conceptual and actionable approach of operationalizing responsibility in speech AI. This governance-centered life-cycle is not presented as a separate contribution from the review itself; rather, it distills and integrates the recurring patterns, gaps, and community priorities identified across the surveyed literature into a cohesive framework for future ASR development.

We define a governance-centered ASR life-cycle as a socio-technical framework that embeds community agency, participatory oversight, and institutional accountability across all stages of ASR system development, from problem framing and data collection to evaluation, deployment, and post-deployment auditing. This framework identifies participatory checkpoints across six stages: (1) problem definition, (2) data sourcing, (3) model training, (4) evaluation, (5) deployment, and (6) post-deployment governance. This framework is not merely a sequencing of technical processes but a normative proposal for reorienting ASR development toward inclusive, just, and reflexive practices. Governance here is not simply oversight after deployment, but a continuous set of ethical, political, and participatory commitments embedded across the ASR pipeline⁴.

A key design imperative, especially for AAE, is acknowledging its fluidity, regional variation, and rapid incorporation of new lexical and phonological features. ASR systems that fail to account for this dynamism risk obsolescence and cultural erasure, since models trained on static datasets will lag behind evolving usage (Farrington 2019; Green 2002b). A governance-centered approach treats this variability not as a modeling problem to be minimized, but as a feature to be preserved through continuous dataset updates, community-led monitoring, and dialect-sensitive evaluation criteria.

We identify three reasons why governance-centered approaches remain underexplored: (1) interdisciplinary disconnects between ML/NLP, HCI, and sociolinguistics hinder holistic system-level thinking; (2) infrastructural limitations, such as proprietary datasets and black-box model architectures, prevent transparency and shared decision-making; and (3) existing incentive structures in academia and industry often prioritize technical novelty over participatory integrity. These challenges make it difficult to operationalize governance, even when its necessity is acknowledged. We find that few studies operationalize participa-

⁴(See the Appendices (A) in extended paper version for visual graph of Governance-Centered ASR Life-cycle Framework.)

Governance-Centered ASR Lifecycle

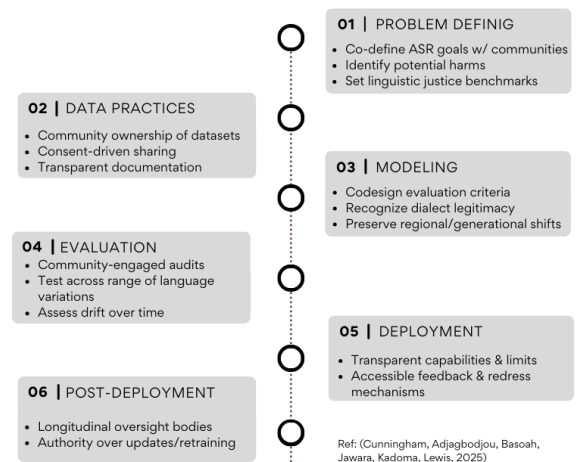


Figure 3: Governance-Centered ASR Framework.

tory governance at scale. Participatory design and annotation studies remain pilot-scale, often lacking institutional continuity, sustainable funding, or decision-making power shared with communities (Sloane et al. 2020). Where community feedback is gathered, it is too often treated as consultation rather than co-governance. Most industry-led audits and bias evaluations are still conducted internally, without transparency, external validation, or community accountability mechanisms.

The value of a governance-centered life-cycle is substantial. It provides a structured lens through which technologists can anticipate and mitigate downstream harms, assess system legitimacy, and design feedback loops that amplify community voice. It aligns with calls from decolonial computing (Irani et al. 2010; Ali 2016), public interest technology (Eubanks 2018), and AI ethics communities that advocate for accountable socio-technical infrastructures. By embedding participation, consent, and oversight at each stage of the ASR life-cycle, this framework offers a roadmap for equitable and culturally responsive speech AI development.

We argue that a framework for a governance-centered life-cycle requires a shift in how ASR systems are built:

- **At the problem definition stage**, governance means defining ASR goals in collaboration with community stakeholders, identifying potential harms, and setting linguistic justice benchmarks before technical work begins.
- **At the data stage**, this means community ownership of speech datasets, consent-driven data sharing, and transparent documentation that centers community goals (Papakriakopoulos and Xiang 2023; Nkemelu et al. 2023).
- **At the modeling stage**, it entails co-designing evaluation criteria, recognizing dialectal legitimacy, and ensuring linguistic variation, including AAE’s regional and generational shifts, is preserved as a design feature rather than

normalized away (Ziems et al. 2022; Prinos, Patwari, and Power 2024).

- **At the evaluation stage**, governance includes community-led audits, testing across regional AAE variants, and assessing model performance over time to catch drift in recognition accuracy.
- **At the deployment stage**, it means establishing transparent communication about system capabilities and limitations, and building accessible user feedback and redress mechanisms.
- **At the post-deployment stage**, governance involves maintaining longitudinal oversight bodies (e.g., community advisory boards, ASR data trusts) with real decision-making authority over updates, retraining, and decommissioning (Blodgett et al. 2022; Ali 2016).

This reframing draws from infrastructure studies (Eubanks 2018), postcolonial computing (Irani et al. 2010; Ali 2016), and recent calls in AI ethics to move from fairness metrics toward democratic accountability (Sloane et al. 2020; Mohamed, Png, and Isaac 2020). It is animated by work showing that community-aligned processes, such as those proposed in Māori ASR initiatives (Brown et al. 2024), Black speech data governance frameworks (Heldreth 2023), and participatory ASR toolkits (Reitmaier et al. 2022, 2023), can produce both culturally competent systems and a deeper trust among impacted communities.

By directly addressing the dynamic, evolving nature of AAE, this approach not only counters the epistemic exploitation and algorithmic erasure identified in critical sociolinguistics (Rosa and Flores 2017), but also ensures that ASR systems remain adaptable to real-world linguistic change. Ultimately, the governance-centered ASR life-cycle advances a justice-oriented vision for speech AI: one where fairness is co-defined, power is redistributed, and community stewardship, not corporate discretion, guides the future of voice-based technologies.

Cross-Disciplinary Blind Spots

Each disciplinary domain contributes partial perspectives. ML/NLP scholarship prioritizes technical rigor but rarely incorporates community-grounded linguistic expertise (Shah, Schwartz, and Hovy 2020). HCI and sociolinguistic studies foreground user experience and cultural specificity but may lack pathways to influence large-scale system development (Bird 2024; Harrington et al. 2022). The result is a fragmented landscape of interventions. Building a governance-centered life-cycle requires epistemic humility (Hanna et al. 2020) and methodological hybridity that bridges these silos. This echoes broader calls in FAccT and AIES scholarship to transcend disciplinary silos when confronting structural bias (Birhane et al. 2022).

Participatory Governance as Ethical Imperative

Participatory design and audit mechanisms offer promising pathways for embedding community accountability (Veale and Brass 2018; Cunningham 2023). However, few studies operationalize community co-governance in the evaluation or deployment of ASR systems. Projects like “Ele-

vating Black Voices” (Heldreth 2023) and Māori-led NLP initiatives (Brown et al. 2024) illustrate models of data sovereignty and cultural stewardship that go beyond inclusion toward structural reorientation. Without such models, the risks of epistemic exploitation and linguistic erasure persist (Blodgett et al. 2022). We urge future scholarship to move from inclusive design to inclusive governance, embedding accountability not only in the system, but in the institutions that build and deploy them.

Study Limitations

In regards to our Methods, we acknowledge that future scoping reviews could benefit from a more systematic documentation of keyword inclusion decisions and filtration pathways, which we identify as a methodological learning from this review process. We also acknowledge that review-writing itself is a form of scholarly gate-keeping (Smith and Dimmick 2010), shaping which works, voices, and epistemologies become canonized in a field where structural inequities already limit whose contributions are visible.

We also recognize the limitations of relying solely on peer-reviewed literature. Our review excludes grey literature, community-authored reports, and archival sources such as Black Web archives. While this decision ensured methodological consistency and comparability across peer-reviewed works, it omits valuable community knowledge and practice-based evidence that may provide further insight into culturally grounded ASR development.

Lastly, we recognize that this review does not include an empirical evaluation of ASR system performance. Future work should pair empirical ASR benchmarking—particularly on African American English and other marginalized varieties—with governance-centered design evaluations to advance recommendations in this review.

Conclusion

This scoping review synthesizes how fairness, bias, and equity are addressed in ASR for AAE and presents a framework grounded in linguistic justice, offering transferable lessons for designing ASR systems inclusive of linguistically diverse communities. Using the PRISMA framework, we synthesized 44 interdisciplinary papers across HCI, ML/NLP, and sociolinguistics, identifying four key domains: (1) evolving understandings of ASR harms; (2) inclusive data practices; (3) methodological and theoretical advances; and (4) practical design and governance approaches.

Our analysis surfaced a persistent gap: while technical interventions abound, few works engage deeply with questions of governance, power, and community agency. In response, we propose a governance-centered ASR life-cycle, a future-oriented framework that re-imagines fairness as a participatory, socio-technical process. We urge future work to advance this direction by co-developing participatory governance models and embedding linguistic justice across the ASR pipeline. Future work should also empirically validate this framework with diverse speech communities by co-developing benchmarks, refining models to community priorities, and evaluating socio-technical impacts.

Ethics Statement

This review synthesizes interdisciplinary scholarship on ASR, fairness, and linguistic inclusion, with a focus on AAE. We acknowledge that this work is not neutral: it involves interpretive choices that shape which perspectives are elevated in a field where Black voices and linguistic diversity are often marginalized.

Our ethical commitments are rooted in our positionality. (Harding 2004), as we advance care for underrepresented communities, epistemic humility across disciplines, and transparency in how we synthesize and critique scholarship. This work is led by Black researchers whose lived experiences and interdisciplinary expertise in HCI and socio-technical AI inform a commitment to addressing systemic anti-Blackness and linguistic erasure.

Acknowledgments

This work was supported by the National Science Foundation Graduate Research Fellowship, the Google Award for Inclusive Research, and the Social Science Research Council Just Tech Fellowship. The first author gratefully acknowledges the commitment of these organizations to advancing research at the intersection of technology equity and social justice.

References

- Ali, S. M. 2016. A brief introduction to decolonial computing. *XRDS*, 22(4): 16–21.
- Allen, B. J. 1996. Feminist standpoint theory: A black woman’s (re) view of organizational socialization. *Communication Studies*, 47(4): 257–271.
- Bender, E. M.; and Friedman, B. 2018. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. *Transactions of the Association for Computational Linguistics*, 6: 587–604.
- Bird, S. 2024. Decolonising Speech and Language Technology. In *Proceedings of the 28th International Conference on Computational Linguistics*, 3504–3519. International Committee on Computational Linguistics.
- Birhane, A.; Isaac, W.; Prabhakaran, V.; Diaz, M.; Elish, M. C.; Gabriel, I.; and Mohamed, S. 2022. Power to the People? Opportunities and Challenges for Participatory AI. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, EAAMO ’22, 1–8. Association for Computing Machinery. ISBN 978-1-4503-9477-2.
- Björgvinsson, E.; Ehn, P.; and Hillgren, P.-A. 2010. Participatory design and “democratizing innovation”. In *Proceedings of the 11th Biennial Participatory Design Conference*, PDC ’10, 41–50. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-0131-2.
- Blodgett, S. L.; Barocas, S.; Daumé III, H.; and Wallach, H. 2020. Language (Technology) is Power: A Critical Survey of “Bias” in NLP. In Jurafsky, D.; Chai, J.; Schluter, N.; and Tetreault, J., eds., *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5454–5476. Online: Association for Computational Linguistics.
- Blodgett, S. L.; Green, L.; and O’Connor, B. 2016. Demographic Dialectal Variation in Social Media: A Case Study of African-American English. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 1119–1130. Association for Computational Linguistics.
- Blodgett, S. L.; Liao, Q. V.; Olteanu, A.; Mihalcea, R.; Muller, M.; Scheuerman, M. K.; Tan, C.; and Yang, Q. 2022. Responsible Language Technologies: Foreseeing and Mitigating Harms. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI EA ’22, 1–3. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-9156-6.
- Bourdieu, P. 1991. *Language and Symbolic Power*. Harvard University Press. ISBN 978-0-674-51041-8. Google-Books-ID: u2ZIGBiJntAC.
- Brewer, R. N.; Harrington, C.; and Heldreth, C. 2023. Envisioning Equitable Speech Technologies for Black Older Adults. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT ’23, 379–388. New York, NY, USA: Association for Computing Machinery. ISBN 9798400701924.
- Brown, P. T.; Wilson, D.; West, K.; Escott, K.-R.; Basabas, K.; Ritchie, B.; Lucas, D.; Taia, I.; Kusabs, N.; and Keegan, T. T. 2024. Māori Algorithmic Sovereignty: Idea, Principles, and Use.
- Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1): 37–46.
- Crawford, K. 2021. *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press. ISBN 9780300209570.
- Crenshaw, K. 1991. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, 43(6): 1241–1299. Publisher: Stanford Law Review.
- Cunningham, J.; Blodgett, S. L.; Madaio, M.; Daumé Iii, H.; Harrington, C.; and Wallach, H. 2024. Understanding the Impacts of Language Technologies’ Performance Disparities on African American Language Speakers. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Findings of the Association for Computational Linguistics ACL 2024*, 12826–12833. Bangkok, Thailand and virtual meeting: Association for Computational Linguistics.
- Cunningham, J. L. 2023. Collaboratively Mitigating Racial Disparities in Automated Speech Recognition and Language Technologies with African American English Speakers: Community-Collaborative and Equity-Centered Approaches Toward Designing Inclusive Natural Language Systems. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA ’23, 1–5. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-9422-2.
- Davidson, T.; Bhattacharya, D.; and Weber, I. 2019. Racial Bias in Hate Speech and Abusive Language Detection Datasets. In *Proceedings of the Third Workshop on Abu-*

- sive Language Online, 25–35. Association for Computational Linguistics.
- Dorn, R. 2019. Dialect-Specific Models for Automatic Speech Recognition of African American Vernacular English. In *Proceedings of the Student Research Workshop Associated with RANLP 2019*, 16–20. INCOMA Ltd.
- Eubanks, V. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press. ISBN 9781250074317.
- Farrington, C. 2019. The evolution of African American English. *Oxford Research Encyclopedia of Linguistics*.
- Forrest, J.; and Wolfram, W. 2019. Regional variation in African American Language. *Oxford Research Encyclopedia of Linguistics*.
- Grabe, E.; Nolan, F.; and Farrar, K. J. 1998. IViE-A comparative transcription system for intonational variation in English. In *Fifth International Conference on Spoken Language Processing*.
- Green, L. J. 2002a. *African American English: A Linguistic Introduction*. Cambridge: Cambridge University Press. ISBN 978-0-521-89138-7.
- Green, L. J. 2002b. *African American English: a linguistic introduction*. Cambridge University Press.
- Groenwold, S.; Ou, L.; Parekh, A.; Honnavalli, S.; Levy, S.; Mirza, D.; and Wang, W. Y. 2020. Investigating African-American Vernacular English in Transformer-Based Text Generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 5877–5883. Association for Computational Linguistics.
- Hanna, A.; Denton, E.; Smart, A.; and Smith-Loud, J. 2020. Towards a critical race methodology in algorithmic fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20*, 501–512. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-6936-7.
- Harding, S. G. 2004. *The feminist standpoint theory reader: Intellectual and political controversies*. Psychology Press.
- Harrington, C. N.; Garg, R.; Woodward, A.; and Williams, D. 2022. Its Kind of Like Code-switching: Black Older Adults Experiences with a Voice Assistant for Health Information Seeking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, 1–15. Association for Computing Machinery. ISBN 978-1-4503-9157-3.
- Heldreth, C. 2023. A partnership with Howard University to improve speech technology for Black voices. Google Blog.
- Hutiri, W. T.; and Ding, A. Y. 2022. Bias in Automated Speaker Recognition. In *2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT '22*, 230–247. Association for Computing Machinery. ISBN 978-1-4503-9352-2.
- Iacobelli, F.; and Cassell, J. 2007. Ethnic Identity and Engagement in Embodied Conversational Agents. In *Proceedings of the 7th international conference on Intelligent Virtual Agents, IVA '07*, 57–63. Springer-Verlag. ISBN 978-3-540-74996-7.
- Irani, L.; Vertesi, J.; Dourish, P.; Philip, K.; and Grinter, R. E. 2010. Postcolonial computing: a lens on design and development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1311–1320. ACM.
- Jo, E. S.; and Gebru, T. 2020. Lessons from archives: strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20*, 306–316. Association for Computing Machinery. ISBN 978-1-4503-6936-7.
- Johnson, A.; Fan, R.; Morris, R.; and Alwan, A. 2022. LPC Augment: an LPC-based ASR Data Augmentation Algorithm for Low and Zero-Resource Children's Dialects. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 8577–8581. ISSN: 2379-190X.
- Joshi, P.; Santy, S.; Budhiraja, A.; Bali, K.; and Choudhury, M. 2020. The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6282–6293. Association for Computational Linguistics.
- Jørgensen, A.; Hovy, D.; and Søgaard, A. 2015. Challenges of studying and processing dialects in social media. In *Proceedings of the Workshop on Noisy User-generated Text*, 9–18. Association for Computational Linguistics.
- Kim, J.; Muhic, J.; Robert, L. P.; and Park, S. Y. 2022. Designing Chatbots with Black Americans with Chronic Conditions: Overcoming Challenges against COVID-19. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, 1–17. Association for Computing Machinery. ISBN 978-1-4503-9157-3.
- King, S. 2020. Language and Race. *Annual Review of Linguistics*, 6(1): 67–83.
- Koenecke, A.; Nam, A.; Lake, E.; Nudell, J.; Quartey, M.; Mengesha, Z.; Toups, C.; Rickford, J. R.; Jurafsky, D.; and Goel, S. 2020. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, 117(14): 7684–7689. Publisher: Proceedings of the National Academy of Sciences.
- Labov, W. 1963. The Social Motivation of a Sound Change. *WORD*, 19(3): 273–309. Publisher: Routledge eprint: <https://doi.org/10.1080/00437956.1963.11659799>.
- Labov, W. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. University of Pennsylvania Press.
- Lanehart, S. L. 2015. *We Be Lovin' Black English*. Routledge.
- Lister, K.; Coughlan, T.; Iniesto, F.; Freear, N.; and Devine, P. 2020. Accessible conversational user interfaces: considerations for design. In *Proceedings of the 17th International Web for All Conference*, 1–11. ACM. ISBN 978-1-4503-7056-1.
- Madaio, M. A.; Stark, L.; Wortman Vaughan, J.; and Wallach, H. 2020. Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, 1–14.

- Manzini, T.; Yao Chong, L.; Black, A. W.; and Tsvetkov, Y. 2019. Black is to Criminal as Caucasian is to Police: Detecting and Removing Multiclass Bias in Word Embeddings. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 615–621. Association for Computational Linguistics.
- Markl, N. 2022. Language variation and algorithmic bias: understanding algorithmic bias in British English automatic speech recognition. In *2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT '22*, 521–534. Association for Computing Machinery. ISBN 978-1-4503-9352-2.
- Markl, N.; and McNulty, S. J. 2022. Language technology practitioners as language managers: arbitrating data bias and predictive bias in ASR. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 6328–6339. European Language Resources Association.
- Mengesha, Z.; Heldreth, C.; Lahav, M.; Sublewski, J.; and Tuennerman, E. 2021. “I don’t Think These Devices are Very Culturally Sensitive.”—Impact of Automated Speech Recognition Errors on African Americans. *Frontiers in Artificial Intelligence*, 4: 1–15.
- Mohamed, S.; Png, M.-T.; and Isaac, W. 2020. Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, 33(4): 659–684.
- Morgan, M. 2002. *Language, Discourse and Power in African American Culture*. Studies in the Social and Cultural Foundations of Language. Cambridge: Cambridge University Press. ISBN 978-0-521-80671-8.
- Ngueajio, M. K.; and Washington, G. 2022. Hey ASR System! Why Aren’t You More Inclusive? Automatic Speech Recognition Systems’ Bias and Proposed Bias Mitigation Techniques. A Literature Review. In *HCI International 2022 – Late Breaking Papers: Interacting with eXtended Reality and Artificial Intelligence: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings*, 421–440. Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-031-21706-7.
- Nkemelu, D.; Shah, H.; Essa, I.; and Best, M. L. 2023. Tackling Hate Speech in Low-resource Languages with Context Experts. In *Proceedings of [Conference/Workshop Name]*.
- Okpala, E.; Cheng, L.; Mbwambo, N.; and Luo, F. 2022. AAEBERT: Debiasing BERT-based Hate Speech Detection Models via Adversarial Learning. In *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1606–1612.
- Papakyriakopoulos, O.; and Xiang, A. 2023. Considerations for Ethical Speech Recognition Datasets. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM '23*, 1287–1288. Association for Computing Machinery. ISBN 978-1-4503-9407-9.
- Poletto, F.; Basile, V.; Sanguinetti, M.; Bosco, C.; and Patti, V. 2021. Resources and benchmark corpora for hate speech detection: a systematic review. *Language Resources and Evaluation*, 55(2): 477–523.
- Poplack, S. 1980. Sometimes I’ll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching1. *Linguistics*, 18(7-8): 581–618. Publisher: De Gruyter Mouton Section: Linguistics.
- Prinos, K.; Patwari, N.; and Power, C. A. 2024. Speaking of accent: A content analysis of accent misconceptions in ASR research. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, 1–10. Rio de Janeiro, Brazil: ACM.
- Racadio, R.; Rose, E. J.; and Kolko, B. E. 2014. Research at the margin: participatory design and community based participatory research. In *Proceedings of the 13th Participatory Design Conference: Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium papers, and Keynote abstracts - Volume 2, PDC '14*, 49–52. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-3214-9.
- Reitmaier, T.; Wallington, E.; Kalarikalayil Raju, D.; Klejch, O.; Pearson, J.; Jones, M.; Bell, P.; and Robinson, S. 2022. Opportunities and Challenges of Automatic Speech Recognition Systems for Low-Resource Language Speakers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, 1–17. Association for Computing Machinery. ISBN 978-1-4503-9157-3.
- Reitmaier, T.; Wallington, E.; Klejch, O.; Markl, N.; Lam-Yee-Mui, L.-M.; Pearson, J.; Jones, M.; Bell, P.; and Robinson, S. 2023. Situating Automatic Speech Recognition Development within Communities of Under-heard Language Speakers. In *CHI '23: ACM CHI Conference on Human Factors in Computing Systems*, 1–17. Hamburg, Germany: ACM.
- Rickford, J. R. 2016. *Raciolinguistics: How Language Shapes Our Ideas about Race*. Oxford University Press. ISBN 978-0-19-062569-6. Google-Books-ID: gVf0DAAAQBAJ.
- Rosa, J.; and Flores, N. 2017. Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society*, 46(5): 621–647.
- Santiago, H.; Martin, J.; Moeller, S.; and Tang, K. 2022. Disambiguation of morpho-syntactic features of African American English – the case of habitual be. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, 70–75. Association for Computational Linguistics.
- Spa, M.; Card, D.; Gabriel, S.; Choi, Y.; and Smith, N. A. 2019. The Risk of Racial Bias in Hate Speech Detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1668–1678. Association for Computational Linguistics.
- Sarı, L.; Hasegawa-Johnson, M.; and Yoo, C. D. 2021. Counterfactually Fair Automatic Speech Recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29: 3515–3525. Conference Name: IEEE/ACM Transactions on Audio, Speech, and Language Processing.
- Shah, D. S.; Schwartz, H. A.; and Hovy, D. 2020. Predictive Biases in Natural Language Processing Models: A Conceptual Framework and Overview. In *Proceedings of the*

- 58th Annual Meeting of the Association for Computational Linguistics, 5248–5264. Association for Computational Linguistics.
- Sloane, M.; Moss, E.; Awomolo, O.; and Forlano, L. 2020. Participation is not a Design Fix for Machine Learning. *arXiv preprint arXiv:2006.07806*, 7.
- Smith, A. D. R.; Ahmed, A. A.; Alvarado Garcia, A.; Dosono, B.; Ogbonnaya-Ogburu, I.; Rankin, Y.; To, A.; and Toyama, K. 2020. What's Race Got To Do With It? Engaging in Race in HCI. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, 1–8. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-6819-3.
- Smith, S. J.; and Dimmick, J. 2010. Gatekeeping Theory: An Evolution. *Journal of Communication*, 60(4): 611–631.
- Sutton, S. J.; Foulkes, P.; Kirk, D.; and Lawson, S. 2019. Voice as a Design Material: Sociophonetic Inspired Design Strategies in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, 1–14. Association for Computing Machinery. ISBN 978-1-4503-5970-2.
- Thomason, S. G. 2021. 14 Managing Sociolinguistic Data with the Corpus of Regional African American Language (CORAAAL). In *The Open Handbook of Linguistic Data Management*, 185–193. MIT Press. ISBN 978-0-262-36607-6. Conference Name: The Open Handbook of Linguistic Data Management.
- Tricco, A. C.; Lillie, E.; Zarin, W.; O'Brien, K. K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M. D. J.; Horsley, T.; Weeks, L.; Hempel, S.; Akl, E. A.; Chang, C.; McGowan, J.; Stewart, L.; Hartling, L.; Aldcroft, A.; Wilson, M. G.; Garrity, C.; Lewin, S.; Godfrey, C. M.; Macdonald, M. T.; Langlois, E. V.; Soares-Weiser, K.; Moriarty, J.; Clifford, T.; Tunçalp, Ö.; and Straus, S. E. 2018. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7): 467–473.
- Veale, M.; and Brass, I. 2018. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.
- Wassink, A. B.; Gansen, C.; and Bartholomew, I. 2022. Uneven success: automatic speech recognition and ethnicity-related dialects. *Speech Communication*, 140: 50–70.
- Wenzel, K.; Devireddy, N.; Davison, C.; and Kaufman, G. 2023. Can Voice Assistants Be Microaggressors? Cross-Race Psychological Responses to Failures of Automatic Speech Recognition.
- Wenzel, K.; and Kaufman, G. 2023. Challenges in Designing Racially Inclusive Language Technologies. *arXiv:2303.13546 [cs]*.
- Weston, D. 2021. Gatekeeping and linguistic capital: A case study of the Cambridge university undergraduate admissions interview. *Journal of Pragmatics*, 176: 137–149.
- Xu, A.; Pathak, E.; Wallace, E.; Gururangan, S.; Sap, M.; and Klein, D. 2021. Detoxifying Language Models Risks Marginalizing Minority Voices. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2390–2397. Association for Computational Linguistics.
- Ziems, C.; Chen, J.; Harris, C.; Anderson, J.; and Yang, D. 2022. VALUE: Understanding Dialect Disparity in NLU. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 3701–3720. Association for Computational Linguistics.