Building Common Ground from the Ground Up: Repair Infrastructure for Human–Agent Collaboration in African Languages

Cathy Holloway, Giulia Barbareschi, Richard Cave, Kathryn Tomanek, Chintan Ghate, Gifty Ayoka, Brian Mwanda, Victoria Austin

UCL Global Disability Innovation Hub / Centre for Digital Language Inclusion

c.holloway@ucl.ac.uk

ABSTRACT

Theories of distributed teamwork portraying LLMs as remote collaborators are frequently constructed around an unexamined assumption: that collaborators share a natural language. For speakers of the vast majority of the world's approximately 7,000 languages, this assumption does not hold—the LLM agents are not merely remote but functionally non-communicative as they do not share the languages of the users they supposedly collaborate with. Drawing on three years of work through the Centre for Digital Language Inclusion (CDLI), which has scaled community-driven speech recognition from one to thirteen African languages, we argue that linguistic asymmetry is the defining yet overlooked barrier to human–agent collaboration for the majority world. Communities who collect, validate, and steward speech data are not passive data sources but essential *third-party collaborators* performing the invisible labor that makes human–agent partnership possible. We extend the workshop's dyadic model into a triadic one and apply a belonging-oriented design framework—grounded in disability justice and trauma-informed practice—to reframe this community labor as *repair infrastructure*: the collective work of building communicative common ground where none yet exists.

CCS Concepts: Human-centered computing → Collaborative and social computing; Accessibility → Accessibility technologies

Keywords: human–agent collaboration, common ground, belonging, repair infrastructure, low-resource languages, speech recognition, articulation work, digital language inclusion, Africa

1 INTRODUCTION

Large Language Models advanced conversational capabilities have shifted the way such technologies are leveraged by users with dynamics moving beyond the ideas of simple tools to be used, towards more interactive dynamics of collaboration [1]. Such conceptualization is grounded in CSCW research into distributed teamwork of humans [2]. It is certainly true that there are many analogies between how collaboration unfolds between users and LLM agents and amongst teams of human collaborators. Namely, during both collaboration between humans and LLM agents as well as among remote human collaborators, communication takes place in digital channels, frequently using text or spoke formats, there is limited to no opportunity to share the same physical space, and iterative interaction is used to establish common grounds for collaboration.

However, the comparison between these two types of collaboration contains an unexamined assumption: that the collaborators share a language. Clark and Brennan's foundational account of common ground [3] takes mutual intelligibility as a starting condition—something collaborators *have* before the process of grounding begins. For remote human collaborators, this is generally true. For the majority of the world's population attempting to interact with LLM agents, it is not.

Of the world's approximately 7,000 languages, the vast majority are "low-resource"—lacking the digitised text and speech corpora that underpin both LLM training and automatic speech recognition [4]. For speakers of these languages, the agent is not a remote collaborator with whom common ground must be *maintained*. It is an entity with whom common ground must be *constructed from scratch*—and that construction requires collective, community-level effort that precedes and enables any individual interaction. The question is not

only how humans and agents collaborate, but *who builds the conditions for that collaboration to become possible, and at what cost*.

In this position paper, we draw on three years of experience through the Centre for Digital Language Inclusion (CDLI) and on recent work in belonging-oriented design [5] to make two arguments. First, that linguistic asymmetry is the *primary* barrier to human–agent collaboration for the majority world, and that this barrier is largely absent from current discourse on agentic AI. Second, that the communities who collectively build the linguistic infrastructure for human–agent interaction are performing a form of *repair work* [6]—constructing the communicative bridge that agent designers failed to build—and that this labor deserves recognition, resourcing, and reciprocity within any account of human–agent collaboration.

2 CONSTRUCTING COMMON GROUND: THE CDLI CASE

The Centre for Digital Language Inclusion (CDLI), based at UCL and supported by Google.org, develops speech recognition technology for African languages [7]. Since its inception, CDLI has scaled from a single language dataset to thirteen languages across multiple countries, including Kenya, Ghana, Uganda, and Rwanda. Critically, the project does not follow a conventional extractive data pipeline. It employs a community-driven methodology in which local entrepreneurs and language communities lead the collection, validation, and curation of speech data [8].

This model was born of necessity. Low-resource languages lack the large-scale digital corpora that enable conventional supervised learning. But CDLI's experience has revealed something more fundamental: the process of building speech datasets is not merely a technical prerequisite for ASR. It is a form of *collaborative infrastructure work* in which communities negotiate linguistic norms, dialectal variation, and the boundaries of what "counts" as their language in digital form. When a community in western Kenya debates whether to include Luhya sub-dialects or treat them as distinct datasets, they are not resolving a labelling problem—they are determining the scope of future human–agent communication. The community is not a passive data source. It is an active agent in constructing the conditions under which human–agent collaboration can occur.

The entrepreneur ecosystem model further illustrates this dynamic. CDLI trains and supports local entrepreneurs across Ghana, Kenya, Uganda, and Rwanda who serve as intermediaries between community language practices and the technical requirements of ASR systems [8]. These entrepreneurs recruit contributors, manage quality assurance, negotiate community consent, and sustain engagement over time. They perform what Schmidt and Bannon [9] termed *articulation work*—the coordinative labour that makes primary work possible but is invisible within the primary activity. In this case, the "primary work" is human–agent interaction. The articulation work is everything a community must do *before* that interaction can happen at all.

To date, CDLI has collected over thirty hours of impaired speech data in Akan through its Ghana pilot, with data collection now extending across Kenya, Uganda, and Rwanda. The project partners with local institutions including the University of Ghana, Talking Tipps Africa, and Senses Hub Kenya, training local entrepreneurs and speech and language therapists to lead community-driven data collection in their own languages. Fine-tuning the open-source Whisper model on this community-collected data yielded a median relative word error rate reduction of 21.7% on the impaired speech test set, demonstrating both the feasibility of the approach and the scale of the performance gap that remains when communities are excluded from training pipelines.

3 FROM COMMON GROUND TO REPAIR INFRASTRUCTURE

3.1 The Absent Baseline

Clark and Brennan [3] identify constraints on grounding—copresence, visibility, audibility, contemporality—that shape the costs of establishing mutual understanding. The workshop's "agent as remote collaborator" analogy maps productively onto this framework: like remote human collaborators, LLM agents operate under reduced copresence, limited visibility, and constrained feedback channels.

But this framework assumes a baseline that precedes all such constraints: that both parties can *produce and comprehend utterances in a shared language*. For low-resource language speakers interacting with current AI

systems, this condition is absent. We are not dealing with *degraded grounding* (the remote collaboration case) but with *absent grounding infrastructure* (a qualitatively different problem). The question shifts from "how do we maintain common ground through lean channels?" to "who builds the common ground, and at what cost?"

3.2 Linguistic Exclusion as Belonging Rupture

The experience of encountering an agent that cannot understand your language is not merely a usability failure. It is what Holloway [5] terms a *belonging rupture*—a moment where the implicit message is not "this tool doesn't work for you" but "you were not considered when this was built." Drawing on Baumeister and Leary's [10] foundational account of belonging as a need for frequent, emotionally positive interactions within stable frameworks of mutual care, and on Allen et al.'s [11] model of belonging in computing as negotiated through membership, competence, and relationships, the belonging-oriented design framework [5] identifies a four-stage cycle through which belonging ruptures and recovers: *Recognition* (awareness of misattunement), *Withdrawal* (protective response), *Reflection* (reframing), and *Repair* (adaptation and advocacy).

This cycle maps precisely onto the experience of low-resource language communities encountering agentic AI:

Recognition. A speaker of Luganda, Twi, or Dholuo encounters a voice assistant, LLM agent, or speech-to-text system and discovers it cannot process their language. The system either fails silently, produces garbled output, or defaults to English. The rupture may be sudden (a failed interaction) or gradual (a growing awareness that these technologies are not built for them). Crucially, the system rarely makes its limitations legible—it does not say "I don't speak your language" but simply fails, implicitly attributing the problem to the user rather than the dataset.

Withdrawal. Users disengage from digital systems and agentic tools that exclude their language. This withdrawal is adaptive, not pathological—a rational response to systems that were not designed for them. But it carries a compounding cost: exclusion from the growing ecosystem of AI-mediated services, information, and economic opportunity. The digital divide deepens precisely where agentic AI promises to bridge it.

Reflection. Communities and advocates begin to reframe the problem. The issue is not that their language is "difficult" or "minor" but that the training data pipeline systematically excluded them. This reframing—from individual deficit to systemic exclusion—mirrors the recognition stage in disability justice: the problem is the design, not the person [12, 13].

Repair. Communities mobilise to build the missing infrastructure. CDLI's community-driven speech data collection is, in belonging-oriented design terms, *repair infrastructure* [5, 6]—the collective labour of constructing common ground where the agent's designers failed to provide it. Entrepreneurs recruit contributors, communities negotiate dialect boundaries, and validated datasets are built through sustained, coordinated effort. This is not simply "data collection." It is a community performing the articulation work [9] that the dominant AI development pipeline never provisioned for.

3.3 The Triadic Model: User, Agent, Community

The rupture–repair cycle reveals a collaboration structure absent from current human–agent interaction frameworks. The end-user who eventually speaks to an ASR-enabled agent is not the only human in the collaboration. Behind them stands a community of contributors who collected the speech data, validated transcriptions, negotiated dialectal boundaries, and trained local entrepreneurs to sustain the data ecosystem. We propose that for low-resource language contexts, human–agent collaboration should be understood as a *triadic* relationship:

(1) the *end-user* who interacts with the agent;

(2) the *agent* whose linguistic capabilities are necessarily bounded; and

(3) the *language community* whose collective repair work constructs the communicative bridge between user and agent.

This triad challenges the dyadic framing that dominates both CSCW's human–human collaboration research and HCI's human–agent interaction research. It also raises urgent questions of *accountability and reciprocity* [14]. If communities bear the articulation cost of making agents linguistically competent, what do they receive in return? How should the benefits of human–agent collaboration be distributed when the collaboration depends on community labour? The parallel with Edie's experience in assistive technology [5]—where a user excluded from speech recognition datasets is told the problem is his speech, not the system's training data—operates at a language-community scale in the CDLI case. Entire languages are excluded, and communities are left to build the infrastructure that developers never provisioned for.

3.4 Design Principles for Linguistically Inclusive Agent Collaboration

The belonging-oriented design framework [5] proposes three principles—Disposition-Aware Agency, Temporal Sovereignty, and Repair Infrastructure—that translate directly into the human–agent collaboration context:

Disposition-Aware Agency. Agents should make their linguistic capabilities and limitations legible. Rather than failing silently when encountering an unsupported language, agents should disclose their competence boundaries: "I have limited capability in Dholuo" rather than producing garbled output. This transparency shifts attribution from user inadequacy to system limitation, supporting the recognition stage of the repair cycle and aligning with the workshop's call for interfaces that make agents' "goals, actions, and limitations transparent to human partners."

Temporal Sovereignty. Communities need control over the pace and terms of their engagement with agent-building processes. CDLI's experience shows that community-driven data collection operates on community rhythms—seasonal availability, trust-building timelines, negotiation of dialectal boundaries—not on product development sprints or funder reporting cycles. Temporal sovereignty means resisting the extractive urgency that treats community data as a resource to be harvested on external timelines.

Repair Infrastructure. This is the principle most directly embodied by CDLI's work. Repair infrastructure treats community adaptation not as an exception but as a core system capability [6]. In the language inclusion context, this means designing agent systems that can incorporate community-contributed speech data, that provide channels for communities to report and correct recognition errors, and that distribute the benefits of improved agent performance back to the communities whose labor made it possible. The workshop's interest in "shared accountability" finds its most challenging test case here: accountability to communities whose collective work underpins an agent's capacity to collaborate.

4 PROVOCATIONS FOR THE WORKSHOP

We offer the following questions to stimulate workshop discussion:

What counts as "collaboration" when the agent cannot understand you? Most human–agent collaboration research assumes functional communication as a starting condition. How does the field's research agenda change when we center the majority of the world's language speakers, for whom this condition is unmet?

Who bears the repair cost of agent collaboration? The belonging-oriented design framework identifies repair as collective, not individual. If communities must mobilize to build the linguistic infrastructure agents need, how should this articulation work be recognized, resourced, and reciprocated? What does "shared accountability" mean when the third party is an entire language community?

Is linguistic exclusion a form of epistemic violence in human–agent systems? When an agent can collaborate fluently in English but cannot process Twi or Luganda, the implicit hierarchy is not merely technical. What frameworks from disability justice and postcolonial computing [12, 15] help us understand this as a question of power, not just data availability?

How does the rupture–repair cycle inform trust calibration? The workshop highlights trust calibration as a research priority. But trust operates differently when a community has invested collective effort in building an agent's linguistic capabilities. Betrayals of that trust—through extractive data practices, abandoned projects, or poor-quality systems that waste community contributions—have communal consequences and compound existing patterns of technological marginalization.

5 CONCLUSION

The vision of LLM agents as genuine collaborators is compelling. But for most of the world's population, the prerequisite for that collaboration—a shared language—does not yet exist. CDLI's work demonstrates both the scale of this challenge and a model for addressing it: one in which communities are recognized not as passive beneficiaries of technological progress but as active repair agents, constructing the linguistic infrastructure that makes human–agent partnership possible.

The belonging-oriented design framework [5] provides a vocabulary for understanding this work—as a cycle of recognition, withdrawal, reflection, and repair operating across individual, community, and systemic scales. The BRIDGE evaluation tool offers a practical means of assessing whether human–agent collaboration systems support or undermine this cycle for linguistically marginalized communities.

We invite the workshop community to extend its theoretical and design imagination beyond the dyadic, linguistically homogeneous cases that currently dominate the field, and to consider what human–agent collaboration looks like when common ground must be built—collectively, painstakingly, and at the community's own cost—from the ground up.

ACKNOWLEDGMENTS

This work was supported by Google.org and the AT2030 program, led by the Global Disability Innovation (GDI) Hub and funded by UK International Development to test 'what works' to improve access to life-changing assistive technology for all. We gratefully acknowledge the community data contributors across Kenya, Ghana, Uganda, and Rwanda whose collective labor underpins the research described here, the CDLI entrepreneur network, and our institutional partners including the University of Ghana, Talking Tipps Africa, and Senses Hub Kenya. The work is part of the Centre of Digital language inclusion (CDLI), which is based at UCL's Global Disability Innovation Hub.

REFERENCES

[1] Yao, B., Chen, J., Chen, C., Wang, A., Li, T. J., & Wang, D. (2025). Through the Lens of Human-Human Collaboration: A Configurable Research Platform for Exploring Human-Agent Collaboration. ArXiv Preprint ArXiv:2509.18008..

[2] Gary M. Olson and Judith S. Olson. 2000. Distance Matters. Human–Computer Interaction 15, 2–3 (September 2000), 139–178.

[3] Herbert H. Clark and Susan E. Brennan. 1991. Grounding in Communication. In Perspectives on Socially Shared Cognition, Lauren B. Resnick, John M. Levine, and Stephanie D. Teasley (Eds.). American Psychological Association, Washington, DC, 127–149.

[4] Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 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 (ACL 2020), 6282–6293.

[5] Catherine Holloway. 2026. Designing for Belonging: The Rupture–Repair Cycle in Human–AI Relationships. In Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26). [Under review]

[6] Steven J. Jackson. 2014. Rethinking Repair. In Media Technologies: Essays on Communication, Materiality, and Society, Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (Eds.). MIT Press, 221–239.

[7] Richard Cave, Catherine Holloway, Gifty Ayoka, Katrin Tomanek, Giulia Barbareschi, and Victoria Austin. 2025. Developing African Language Models for Atypical Speech. [In review].

[8] Sumaya Ahmed Salihs, Isaac Wiafe, Jamal-Deen Abdulai, Elikem Doe Atsakpo, Gifty Ayoka, Richard Cave, Akon Obu Ekpezu, Catherine Holloway, Katrin Tomanek, and Fiifi Baffoe Payin Winful. 2025. A Cookbook for Community-driven Data Collection of Impaired Speech in Low-Resource Languages. In Proceedings of Interspeech 2025, 4623–4627.

[9] Kjeld Schmidt and Liam Bannon. 1992. Taking CSCW Seriously: Supporting Articulation Work. Computer Supported Cooperative Work 1, 1–2 (March 1992), 7–40.

[10] Roy F. Baumeister and Mark R. Leary. 1995. The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. Psychological Bulletin 117, 3 (1995), 497–529.

[11] Sarah Allen, Sayam Ahmed, John McCarthy, and Christopher Frauenberger. 2022. Belonging in HCI: Towards a Conceptual Framework. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22), 1–15.

[12] Cynthia L. Bennett and Daniela K. Rosner. 2019. The Promise of Empathy: Design, Disability, and Knowing the "Other." In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), 1–13.

[13] Katta Spiel, Kathrin Gerling, Cynthia L. Bennett, Emeline Brulé, Rua M. Williams, Jennifer Rode, and Jennifer Mankoff. 2020. Nothing About Us Without Us: Investigating the Role of Critical Disability Studies in HCI. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, 1–8.

[14] Batya Friedman, Peter H. Kahn Jr., and Alan Borning. 2008. Value Sensitive Design and Information Systems. In The Handbook of Information and Computer Ethics, Kenneth Einar Himma and Herman T. Tavani (Eds.). Wiley, 69–101.

[15] Lilly Irani, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E. Grinter. 2010. Postcolonial Computing: A Lens on Design and Development. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10), 1311–1320.

[16] Janet X. Chen, Allison McDonald, Yixin Zou, Emily Tseng, Kevin A. Roundy, Acar Tamersoy, Florian Schaub, Thomas Ristenpart, and Nicola Dell. 2022. Trauma-Informed Computing: Towards Safer Technology Experiences for All. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22), 1–20.

[17] Gifty Ayoka, Giulia Barbareschi, Richard Cave, and Catherine Holloway. 2024. Enhancing Communication Equity: Evaluation of an Automated Speech Recognition Application in Ghana. In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), 1–16.

[18] Victoria Austin, Gifty Ayoka, Giulia Barbareschi, Richard Cave, and Catherine Holloway. 2025. Exploring the Usability of Gaze-Based Mobile Communication in Ghana. In Proceedings of the 27th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '25), 1–13.