



Enhancing Communication Equity: evaluation of an automated speech recognition application in Ghana

Gifty Ayoka*
Talking Tipps
Accra, Ghana
giphty2001@gmail.com

Richard Cave*
University College London
London, United Kingdom
richard.cave.17@ucl.ac.uk

Giulia Barbareschi*
Keio Graduate School of Media Design
Yokohama, Japan
barbareschi@kmd.keio.ac.jp

Catherine Holloway*
Global Disability Innovation Hub, UCL
London, United Kingdom
c.holloway@ucl.ac.uk

ABSTRACT

In Ghana people who struggle to articulate speech as a result of different conditions experience barriers in interacting with others due to difficulties in being understood. Automatic speech recognition software can be used to help listeners understand people with communication difficulties. However, studies have not looked at the practical feasibility of these technologies beyond the Global North. We present a novel user study examining the introduction of one such technology, Google Project Relate, to Ghana. This freely available mobile application can create personalised speech recognition models in English for non-standard speech to support communication. Our user study spans the training of local speech and language therapists and 20 people with communication difficulties. We utilise the Technology Amplification Theory to contribute insights on the need for technological adaptations, awareness and support to reduce differential gaps of access, capacity and motivation to expand the reach of these technologies rather than exacerbating inequalities.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in accessibility**.

KEYWORDS

Speech, Disability, Automatic Speech Recognition, Technology Amplification Theory, Ghana, Global South, Accessibility

ACM Reference Format:

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)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3613904.3641903>

*All authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '24, May 11–16, 2024, Honolulu, HI, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0330-0/24/05
<https://doi.org/10.1145/3613904.3641903>

1 INTRODUCTION

Although the global prevalence of communication disabilities is unknown, studies have estimated that as many as 28-49% of people with disabilities worldwide experience difficulties with communication at some point in their lives [39]. Impaired speech is often associated with severe stigma, and people with communication difficulties are amongst the most marginalised groups in society due to existing barriers and discriminatory attitudes [50, 53, 93]. To avoid failed exchanges and misunderstandings, a person with impaired speech may choose to interact only when sharing essential information, choose only to speak with a familiar communication partner, or let others speak on their behalf [23]. A failed social exchange can carry the message of inferiority [30], resulting in reduced participation [79]. Talking remains the preferred medium of communication for many despite the difficulties they might face, as it represents a powerful medium of identity [20], which people can leverage to communicate mood, humour, geographical, social and educational background, health status, gender - as well as the content of the message [61]. If talking becomes difficult to understand, individual and social identity can be negatively affected, increasing the risk of social withdrawal [102].

In Ghana and other West African countries, these barriers are often more pronounced as a result of compounding factors that range from the lack of Speech and Language Therapy (SLT) services and poor availability of assistive technologies to support communication [23, 38, 59, 101], to the stigmatising cultural beliefs that label disability as a curse [98, 101]. Within this context, the expanding technological infrastructure and the increasing penetration rate of mobile phones across all population segments offer a two-fold opportunity. First, a viable opportunity to expand the training and long-term professional development of Speech and Language Therapists (SLTs) in Ghana [38], Second, mobile phones also represent a vital asset to support people who experience difficulties with communication in their everyday lives [12, 13, 33, 60, 67, 76, 81].

Recent developments in the creation of bespoke language models have expanded the possibility of using Automatic Speech Recognition (ASR) software for people with dysarthria, the collective term for a group of neurologic speech disorders linked to muscular dysfunctions, or other conditions that affect the ability to articulate speech [72, 73, 86, 103]. These technologies do not alter how the disabled person speaks but can help listeners better understand what is said by repeating or transcribing words and sentences in

real time to facilitate communication [16, 17, 58]. An example is Project Relate, an English-based ASR application freely available for mobile phones featuring Android 8 and above. It specifically targets individuals with ‘non-standard’ speech, described as speech that differs from the accepted and recognisable speech of adults in a particular language [1]. The application requires a minimum of 500 samples of an individual’s speech, collected by recording several pre-set phrases, to build a customised speech model, which can then be used to produce automated real-time transcription, facilitate interaction with Google Assistant, enable voice typing for SMS and other functions, as well as speech repetition using a synthesised easy-to-understand voice [2].

The Technology Amplification theory [87] illustrates how technology does not deliver additive or transformative benefits without adequate support and infrastructure but amplifies current trends and social inequalities. The mechanisms of amplification revolve around three dimensions: access, capacity, and motivation. Disregarding these differential aspects means that innovative technologies such as Project Relate may only positively impact a small minority of potential users with communication difficulties, while leaving the majority of others in even more marginalised positions.

Although Project Relate officially became available in Ghana in 2022 [1], we noticed very limited awareness of it within Disabled People’s Organisations (DPOs) and SLTs organisations in Ghana. Furthermore, as the application had been developed in the US and it is available only in English, it was unclear to what extent the different contexts could affect its use. Finally, as is the case for most digital products and assistive technologies, users are likely to require specific resources, competencies and support to integrate Project Relate into their life and benefit from its use. However, to date, there is limited guidance available for potential users of Project Relate, such as training, technical support, clear discussion on limitations and potential issues, and advice about navigating barriers or troubleshooting potential issues. It was unclear whether any of these factors impacted the use and usefulness of Project Relate to Ghanaian users.

To identify the mechanisms that determine differences in access, capacity and motivation among users of Project Relate in Ghana, we conducted a 6-week study involving 10 SLTs and 20 adults with communication difficulties. In line with other HCI studies examining the use of mobile phones amongst marginalised populations, we leveraged the lens of the Technology Amplification Theory [41, 87, 91, 96] to analyse data collected during training observations, semi-structured interviews and 4-weeks of self-reported accounts from participants using a photovoice approach.

Our results show differential access to the application and its features is determined by a variety of factors, including the severity of the dysarthria and the presence of other functional impairments, the type of phone, the availability of data, and the language of the user as well as their conversation partner. Differential capacity is affected by the person’s literacy, their ability to create and record customised sentences, which strengthen the language model, making Project Relate more effective in everyday life, and the stigma preventing people from interacting and communicating with others beyond their immediate circle. Finally, differential motivation depends on individuals’ specific life circumstances, which determine

various use-cases. As Project Relate is meant to be used in conversational settings with another person, we highlight how the user’s motivation alone is insufficient; their conversational partners must also accept this new form of interaction.

Based on these results, we provide three key recommendations for the development and deployment of assistive and accessible technologies for communication in Ghana and potentially other contexts in the Global South: 1) Understanding the contextual nature of language, not only in relation to different national languages, but also vocabulary, expressions at an individual and social level; 2) Considering stakeholders beyond the users including their support structure within and beyond the family, SLTs as well as strangers to help normalise perception around technology-mediated communication; 3) Acknowledging strengths and limitations of ASR to understand in which situations they can be beneficial and when other strategies are preferable.

In summary, our study makes the following contributions to HCI:

- A novel study examining the experiences of people with communication difficulties in Ghana with mobile phone-based ASR technologies for non-standard speech
- An analysis of the factors determining differential access, capacity, and motivation in the use of mobile-phone based ASR technology for non-standard speech which can exacerbate existing social inequalities
- A series of recommendations targeting HCI researchers and developers of ASR and other communication technologies specifically targeting users with communication difficulties living in the Global South

2 RELATED WORKS

2.1 Communication Disabilities in Ghana

Ghana is a country of 29.5 million people [95], in common with many countries, there are discrepancies in the estimates for the number of people with disabilities. A household survey in 2010 estimated 3% of the population had a disability [28]; a year later, the World Bank’s Report on Disability reported a figure of between 7-12% [95] and showed a disparity in gender with higher rates for women (10.6% compared to men (6.2%) [95]. Others who focus only on communication difficulties estimate that 6-7% of the population have a communication difficulty [34], equating to approximately 2 million people. Difficulties in articulating speech can emerge due to a variety of congenital, acquired and progressive conditions, including Cleft Palate, Cerebral Palsy, Parkinson’s Disease, head and neck cancer, or ALS. Speech change can occur due to trauma, surgery and even infection [42, 53, 64, 75].

The number of SLTs supporting the communication training and needs of people with communication difficulties also varies from 10 in 2013 to 5 in 2017 [24, 57]. Regardless of which of the two estimates is the correct one, the number is incredibly low, leaving each SLT with at least 200,000 clients needing services and leaving the vast majority of people with communication needs without the support they require. Currently, the SLT services within Ghana sit as a paid-for service not included in the national medical insurance system. Moreover, there is limited (if any) knowledge of communication disability within the community health teams [38, 98]. This is

not uncommon in West Africa or the Global South, more generally, as governments struggle to navigate competing priorities for development within a low or middle-income country. Although Ghana does not yet have a national communication disability rehabilitation service, Wickenden 2013 [94] developed a series of indicators indicating when a country is ready to develop such a service. Ghana is in a position to consider streamlining SLT services, as seen from the development of Disability Rights legislation and general economic growth, and has had a stable democracy for over 25 years [98]. Ghana has also started rolling national training programmes for SLTs, and Ghanaian SLTs have critically reflected on the role of SLTs within an LMIC (or Majority World) context [43, 99].

In 2013, Crowley and colleagues reflected that there were at the time only 10 SLTs in Ghana, and nearly all had been trained overseas. However, they also point to a confluence of factors that showed hope and progress for further momentum in developing the nascent sector of support for people with communication difficulties [24]. Recent work highlighted the ‘emerging’ nature of the SLT workforce in Ghana and the need to educate this first generation of SLTs, especially in using digital tools to support communication, which is well suited to Ghana’s growing and robust digital infrastructure [38]. Previous studies on mobile phone usage amongst people with disabilities in Ghana have shown not only the potential of these tools to increase access to services and information [7, 77], but also a widespread level of digital fluency that can help people take advantage of them [5, 68]. In the next section, we explore in more detail how technology can support people with communication difficulties, specifically examining the possibilities and limitations of ASR for non-standard speech.

2.2 AT for Non-standard Speech Recognition

When speech is difficult to understand, Augmentative and alternative communication (AAC) is sometimes used to support the individual [10]. AAC is a set of tools and strategies a person uses to solve everyday communicative challenges [15]. However, communication using AAC is often slow, around 8–10 words per minute (wpm), increasing to 20 wpm when using word prediction, compared to the average speaking rate of 125 to 185 wpm [46]. As a result of these differences, people may choose not to communicate beyond their essential basic needs [92], or to communicate at all, based on the perceived importance and value of the message and the time available to communicate [23]. Even for those who face significant challenges in being understood, using speech to communicate may remain the first preference in most situations [84], regardless of speaking being effortful and misunderstandings being frequent [14].

In Ghana, the options for AAC are limited due to cost and availability of support [65, 101]. In a similar fashion to what has been observed in other Sub-Saharan countries, due to the existing limitations, the use of AAC devices that substitute speech as a primary form of communication is relatively uncommon in Ghana, even when speech is significantly impaired [25, 40, 100].

An alternative to AACs requiring people to type words or sentences using letters or icons, are ASR technologies. These technologies do not require alternative forms of input but transform

one’s speech into text or provide repetition using an easier-to-understand synthesised voice, facilitating communication with the listener [72, 73]. Increasing the speed of AAC using ASR technology to support understanding of dysarthric speech may contribute to a reduction of the physical, cognitive and emotional effort required by an individual with dysarthric speech – factors that drive the potential for successful communication [23].

ASR accuracy for commercial systems is now as high as 95% for many speakers with unimpaired speech [85], improving significantly over recent years due to the increased computational power of deep learning systems and the availability of large training datasets [32]. However, it may still make too many word recognition errors to benefit people with dysarthria [11]. Very high word error rates were measured for people with significant dysarthria using three commercially available ASR systems (IBM Watson, Google Cloud, and Microsoft Azure Bing) [27], or individuals with severely dysarthric speech, the Word Error Rate (defined as the ratio of errors in a transcript in relation to the total number of words spoken) were between 78–89%, resulting in a very low percentage of correctly transcribed sentences – between 0–1.2% [27].

In recent years, the development of novel algorithmic approaches to speech modelling and recognition has expanded the ability to adapt ASRs to people who are classified as having non-standard speech, including non-native speakers [70], those with different accents [54], people from racial minorities [45], and individuals with dysarthria [49, 73, 80, 86]. However, most existing ASR systems for dysarthric speech are confined to research applications which are unavailable to the public [21, 49, 86], targeting only individuals who have mild dysarthria [56], or focusing on tasks that involve a very limited vocabulary [44].

The ASR application called Project Relate, which was released by Google in 2022 and is currently available only in English, promised to tackle many of these barriers [2]. Built on the large bank of over 1 million utterances by more than 1000 speakers collected as part of the Euphonia project [3], Project Relate enables individuals with non-standard speech to train speech recognition models specifically on their own speech patterns by recording a minimum of 500 pre-set phrases using the only initially available function called “Record” [2]. Users can then leverage their customised voice model through the following four additional functions available in the app:

- (1) **Listen:** Live speech transcription with no time limit. Resulting transcripts can be modified, copied and read out loud. The function can be used offline.
- (2) **Repeat:** Automated playback of short sentences, maximum 10 seconds, using an easy-to-understand synthesised voice. The function needs internet connectivity
- (3) **Assistant:** Supports interactions with the Google Assistant system with personalised ASR and removing existing time constraints. The function needs internet connectivity
- (4) **Keyboard:** Speech to text alternative to standard keyboard integrated with other applications (messaging, email, browser). The function can be used offline.

Although not listed as one of the main five functions (Record, Listen, Repeat, Assistant, and Keyboard), special attention needs to be given to the Custom card, a sub-function accessible under Record,

which allows users to manually add phrases to include in the language model to be created. This is essential to make Relate better understand the key nouns and phrases that a person uses in daily conversation, which might not be part of generalised vocabulary (i.e. names of people, addresses, locations...). Screenshots from each of the functions are shown in Figure 1 below.

A longitudinal study of three people with ALS in the UK reported varying experiences of using Relate to support being understood in everyday communication. All participants reported the significant impact of voice change and loss to their sense of identity and declining sense of ability and competence to participate in social and work activities. Relate was seen as something that could help them be understood by unfamiliar listeners and to a way to participate more in life, but was not as accurate as expected in captioning speech, particularly recognition of proper nouns. All felt the app was difficult to understand at first and required support from technically savvy others. The participants often chose to use Relate in specific contexts, frequently with unfamiliar listeners in transactional exchanges. For other situations they relied on their natural speech, familiarity of conversation partner, combined with slower speed AAC such as a pen and paper when needed. None reported a concern with availability of data or a smartphone to use Relate [in review].

Unfortunately, to date, none of the studies that have looked at the practical implementation of ASR technologies for people with communication difficulties has ever been carried out in Ghana nor, to the best of our knowledge, anywhere in the Global South. Previous research from the HCI, Accessibility and ICT4D communities has examined how infrastructural, societal and personal factors dramatically change how disabled people in the Global South access and leverage mobile phones as assistive technologies, meaning that evidence from research carried out in the US or Europe does not reflect the realities of other geographical contexts and it is in fact more likely to perpetuate misguided colonial labels [4, 12, 13, 41].

2.3 Technology Amplification Theory

In 2011 Kentaro Toyama formulated the Technology Amplification Theory, arguing that “*information and communication technologies have a multiplicative, and not additive, effect on human and institutional intent and capability*” [87]. Technology alone does not create resources or intent where they do not already exist, but it can augment them when they are present. In itself technology is neither a positive nor a negative force; it is simply a tool to be used. One of Toyama’s key observations is that technology is more likely to succeed when it is amplifying already successful interventions that build on existing capacity and positive intent [87]. If the landscape of SLTs services in Ghana is changing, opening up opportunities for people who have difficulties articulating speech, could ASR tools such as Google Relate amplify these possibilities?

One important factor to remember when examining this question is to acknowledge that technology requires infrastructural support at several levels to deliver beneficial effects. Some examples of required support are access to electricity or connectivity, availability of skilled workers to train users and support them, presence of material resources needed for repair and upgrades, as well as the social willingness to accept the technology at a broader level [52, 71,

97]. Where one or more of these elements are missing, technology simply increases existing societal gaps, widening inequalities and leading to further marginalisation [4, 66]. These mechanisms of amplification can be articulated around three axes:

- **Differential Access:** The simple fact that a technology exists does not mean one can access it. Devices require money to be purchased and operated; software is only available in a handful of languages generally spoken by people of higher socio-economic status [4, 13, 74]. Accessibility, or inaccessibility, of physical and digital systems also impacts who is able to access what and which contexts. Barbareschi et al. [12] showed that button phones, more commonly available and preferred by visually impaired users in Kenya, did not include features that allowed them to use the phonebook or listen to text messages, leading participants to classify these services as “*not for blind people*”.
- **Differential Capacity:** Although a person might be able to access a particular technology, the extent to which they could benefit from it would be influenced by a series of factors that influence their personal capacities. Kenyan rural women taking part in the study by Wyche et al 2016 [96] spoke of difficulties performing top-up operations due to low digital literacy, poor eyesight and reduced lighting condition in their homes. Similarly, disabled people interviewed by Jones & Pal 2015 highlighted how acquiring digital skills to operate ICT technologies required significant time, especially for those who had received limited education, which was not necessarily available to those who had no financial stability [41]. Capacity does not only refer to the individual alone, but encompasses their social network [12, 13, 90]. Rural Indian women who had the tools and expertise to engage in crowd work using their mobile devices, struggled to feel valued and thrive as their efforts were easily dismissed by their families [90].
- **Differential Motivation:** Finally, when one is able to access and capable of operating technology, the ultimate use would be decided by a person’s specific motivations. The ICT4D agenda is usually dictated by founders keen on measuring the value of specific programs to justify their investments [4, 36, 78]. Yet, users’ wishes regarding technology are naturally miscellaneous and encompass more “frivolously labelled desires such as play and entertainment [78, 87]. But instead of dispensing easy moral judgments from a position of privilege, funders, researchers and practitioners should consider what might enhance or dampen one’s motivation [4, 41]. Recent work with local farmers in rural India shows that relative advantage and self-efficacy play a huge role in people’s desire to leverage technology for developmental purposes [31]. Simply speaking, if due to systemic barriers, there is no expectation of a beneficial outcome due to systemic barriers, there is no reason to invest effort. Furthermore, while users’ motivations are often scrutinised, overlooking the potential of amplifying negative motivation from wider societal structures is a much more common pitfall. Nova et al. 2019 [62] showed how women and those from more disadvantaged backgrounds were more likely to be subjected

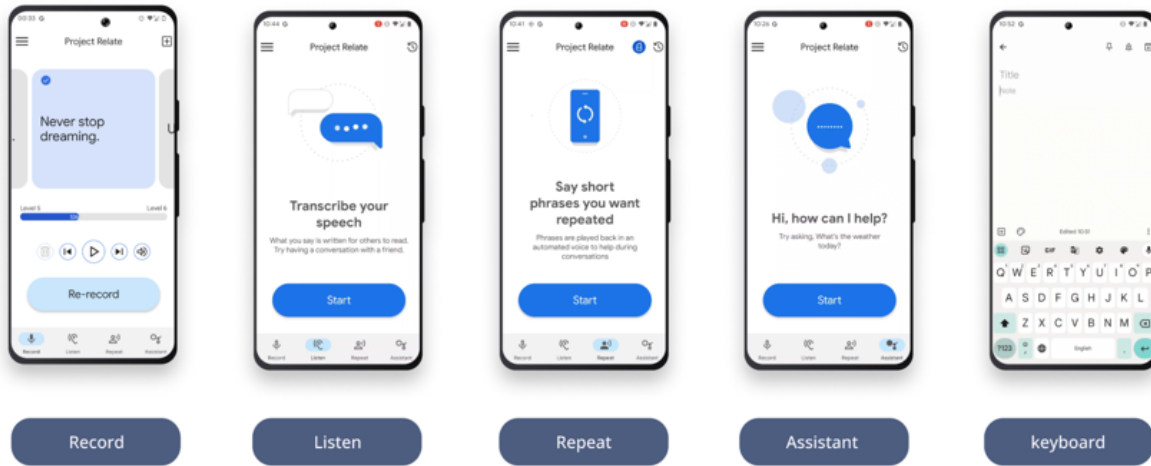


Figure 1: Series of screenshots showing the different view associated with Relate’s functions

to sexual harassment or abuse on anonymous SNS, which offered an easy cover to abusers and others with malicious intent.

Our goal was to evaluate to what extent ASR applications like Google Relate can provide additive benefits to the changing landscape of SLT services and resources for people with communication difficulties in Ghana. To do this, we sought to identify the existing mechanisms of differential access, capacity and motivation that hinder users’ ability to take advantage of the application and suggest strategies for countering them. In the following section, we outline our own positionality, which motivated and drove this work, as well as the detailed approach of our research.

3 METHODS

3.1 Positionality and Collaboration with local SLTs

Our values and positionalities have greatly shaped our approach and commitment to our research. We present our positionalities as a collective as we strongly believe that throughout this process, we have worked as a collective with no hierarchical order, reflected in our choice to list all authors alphabetically while specifying that we all contributed equally to this research according to our skills and expertise. Together, we share various backgrounds, some of us are from the Global South and others from the Global North, and have different experiences working and conducting research across both geographies. Some of us have lived experience of disability, including difficulty articulating speech. Others have indirect experiences as parents, friends of disabled people, or both. Some of us have an academic background with expertise ranging from Computer Science, HCI, Assistive Technology and Speech Therapy, whereas others identify as clinicians or have intersectional professional identities. Our choice of conducting this research in Ghana was motivated by both the personal connection of one of

the authors, who is a Ghanaian national living and working in the country, as well as the fact that Ghana was the first country in the Global South - and the only one in Africa at the time of writing - in which the Google Relate application had been released.

Many of these aspects overlapped with the positions of our participants, which allowed us to foster a sense of connection. However, the experiences, desires, hopes and frustrations we present in the results section are theirs alone. Above anything else, our positionalities are what drove us to adopt a stance that focused on people’s aspirations and sees technology, including Google Relate, as a potential utilitarian tool. Our goal in engaging in this project was not to “*make the application work for people*” but to find ways to promote Equity in communication for people in Ghana who struggle to articulate speech. As part of this, we wanted to understand when the Google Relate application could be leveraged to facilitate communication and what technological and systemic changes might be needed for it to deliver maximal value to as many users as possible.

Our professional, academic, and personal experiences had taught us before that if we wanted our research to deliver actual change and ensure that participants would be supported beyond the lifespan of the project, we had to start by building capacity in situ. To this end, we leveraged our connections to recruit 10 SLTs operating in different parts of the country who had experiences working with a variety of clients with dysarthria and other conditions which affect the ability to articulate speech and be understood in everyday conversations. Before the start of the research, SLTs took part in a training session to illustrate the features of the Google Relate application, explain and trial the process of recording preset phrases and creating new ones using the Custom card function, discuss potential challenges that participants might face when using the application, identify existing design flaws specific to the Ghanaian context, and develop strategies to support future clients interested in using the application. Following the training, the SLTs shadowed research team members during onboarding sessions with

participants, slowly taking over participants' training to ensure they would feel prepared to assist new clients by themselves in the future. Finally, one SLT was included in the communication and support channel created on WhatsApp for each participant to provide clients with a point of contact close to them, which could also provide in-person assistance when required. A separate WhatsApp channel to connect all SLTs with the research team and discuss important matters as they arose was created. Since the start of the project in July 2023, we have continued to use these channels to remain connected not just with participants but also SLTs, celebrating success, troubleshooting difficulties and leveraging our relative privilege to help escalate criticism and call for action with the Google Team in charge of Project Relate.

3.2 Participants

We recruited participants with the support of local SLTs and organisations for people with disabilities in Ghana. SLTs and local organisations identified participants as being difficult to understand by unfamiliar conversation partners due to significant dysarthria, dysphonia or a stutter. To qualify for participation, individuals had to be above the age of 18, able to provide informed consent. As Google Relate is currently available only in English, our biggest restriction when recruiting was the requirement for them to be able to understand and communicate in English in everyday situations, which is not the first language of many Ghanaians [6]. There was no restriction on aetiology, but participants were only enrolled in the study if speech was their primary communication modality (as we do not consider Sign Language users to have communication difficulties in their preferred modalities). Before this research, none of the participants were aware of or had used Google Relate. All participants had access to a mobile phone, but several did not own a smartphone with the required specifications to download and use the Google Relate application. If participants' mobile phones did not meet the minimum specification (Android version 8 or above and at least 1GB of RAM), we provided them with a Samsung AU4s, which they could retain beyond the end of the project. If participants already possessed a smartphone meeting the required characteristics, we provided the equivalent compensation of approximately 120 USD. In our exchanges with the ethics committee at University College London we identified the potential risk of coercing participants to join the research out of interest in obtaining a phone or receiving substantial compensation. To mitigate this, we leveraged the connection and existing relationships of trust that participants had with local organisations and SLTs. Participants were explained how joining the study was voluntary and they could withdraw at any time and retain their phone or the compensation provided as an alternative. Their participation in the study would also not affect any SLT services they were already receiving or would request in the future. However, prolonged engagement with the research team could mean that we could provide support with the application if they needed it and help get their voices heard as we planned to lobby for requested changes with the Google Relate Team.

3.3 Procedure and Data Collection

Onboarding sessions with participants were carried out at the office of the Talking Tippi Africa Foundation in Accra or at participants

homes, depending on their preference. Onboarding sessions were conducted by at least two members of the research team and one of the local SLTs. Participants were invited to bring along a family member or a support person if they wished to do so. Informed consent procedures also included discussions with participants about differences in data collection processes and data storage that existed between the use of the Google Relate application and participation in research. We explained to participants how recordings from interviews and videos and pictures from ethnographic observations captured by the team could be deleted if they requested it, as we had full control over them. However, voice samples taken from the Google Relate application could only be deleted if the participants reached out to Google directly. The privacy notice of the application states that voice samples would only be leveraged for creating the personalised speech recognition model and are not available to third parties nor collated across different users. We explained the privacy policy included in the Google Play store to participants and show them how they could contact the Google team directly from the application "Help" section.

After obtaining informed consent from participants and setting up their mobile phone and a Google account, if needed, we assisted participants to download the Google Relate application from the Play Store. Participants were then supported in recording the sentences needed for the application to create the bespoke voice model and shown how to create custom cards to ensure that the application could understand names and any other word that was important to them in everyday conversations. We explained to participants that to maximise the model's ability to recognise specific words they would have to create three custom cards for each word, changing its position in the phrase from beginning, middle, and end in the following format.

- XXX is my name
- The person XXX is me
- My name is XXXX

Participants were shown this pattern for multiple chosen words and given the opportunity to practice until they felt comfortable they could create Custom cards independently in the future. Some participants were able to complete the recording of all the 500 required phrases during the onboarding session, whereas others due to fatigue or time constraints preferred to continue recording sentences at their own pace independently or with the support of family members, friends or SLTs. A member of the research team took pictures, field notes, short videos and voice recording to document these sessions.

The Google Relate application requires between 24 and 72 hours to create the custom voice model for each user, meaning that even those participants who completed the 500 recordings were not able to test their application immediately. To show the various functionalities, a member of the research team used a mobile phone with the pre-loaded Google Relate application. Participants were able to practice how to enable and use the various functions, albeit using an existing voice model which was highly prone to errors as it was not specific to their speech patterns. Finally, we conducted semi-structured interviews with participants to understand their existing communication difficulties in daily life, their aspirations for a more equitable and accessible communication in various aspects

of life, the expectations they had for the application, difficulties encountered during the onboarding session, and any other feedback about the session or the application.

At the end of the onboarding session we set up the WhatsApp group including participants, the two members of the research team and the SLT who supported the onboarding session and, if requested by participants, their desired support person. Participants were explained that we would use this channel to check in with them periodically and ask for feedback. In turn, they could choose to share any comments, meaningful episodes, as well as request support if needed. Written messages, voice messages, pictures and videos shared by participants were retained for analysis only after obtaining explicit consent. Participants were explained that they could leave the group at any time if they wished to do so.

Situated ethnographic observations to understand how participants used the application in everyday life were carried out with seven participants after they had downloaded their customised voice model. Participants were asked to choose one situation from their everyday life in which they wanted to try to use the application and which they were happy for us to observe. These interactions were video recorded with the consent of participants and the conversation partner when relevant.

3.4 Data Analysis

Our data corpus consists of notes, written messages, pictures, videos and transcribed recordings of semi-structured interviews and voice messages from participants. We analysed data collaboratively using reflexive thematic analysis with a bottom-up approach, with members of the research team reflecting on codes individually and discussing them with the other members during debriefing discussions after each session and a total of 6 group meetings [18, 19]. To articulate insights in a cohesive and easy to understand manner, we decided to leverage the Technology Amplification Theory and conducted 3 group meetings to collaboratively decide how to structure our findings presented in the following section.

We took note of the WER score generated by Relate for each individual's personal ASR model. WER is the most commonly used metric for evaluating speech recognition performance [55]. WER is utilised in Relate in the Profile tab to set user expectations by allocating the personal WER score across 'speech recognition readiness' range from low to high. Relate WER is also detailed in the app's console tab. However, when calculating WER, all words are considered equally important, and all errors (substitutions, deletions, insertions) have equal weight. WER does not consider whether some words may be more important to the meaning of the message and the impact of word errors may be also dependent upon the specific application in which ASR is being used [29].

Additionally, Relate WER is calculated using a sample of standard set phrases and therefore less likely to be phrases a person uses in daily life [51]. The Relate-calculated WER and actual WER scores when Relate is used are likely to be different, indeed varying according to the words a person uses, context and location.

A measure for ASR should have ecological validity that realistically simulates how ASR output would actually be used and how

useful that output would be [29]. People that use ASR for transcription reported less concern about measuring word-for-word accuracy between the spoken message and its transcript but whether the transcript produced by ASR captured the meaning of the spoken message [22]. It is quite possible that a person using a Relate model with a higher WER may find a much or more functional use than an individual with a lower WER model, and vice versa. For these reasons our focus is on how people view the usefulness of Relate in daily life.

4 RESULTS

Twenty people participated in the study. Thirteen participants have dysarthria as the main condition affecting speech, two participants are living with aphasia and dysarthria as a result of stroke, one person has a stutter, whereas four participants have dysphonia - a change to the quality of voice - due to a laryngectomy procedure as a result of head and neck cancer and use a Tracheo-oesophageal puncture (TEP) valve to speak.

Eighteen participants completed the minimum 500 phrases required to create a personalised speech recognition model in Relate. One participant could not record enough phrases even with support from their SLT due to the effects of aphasia due to stroke. The other has a laryngectomy and experienced significant fatigue repeating the phrases using his TEP valve.

Two research team members are certified speech and language therapists with more than 22 years of specialist experience working with people living with voice and speech difficulties. We assessed the severity of impairment for each participant based on our conversations with them, considering articulatory, phonatory, resonatory, respiratory, and prosodic deficits - factors that can significantly limit communicative capacity [48]. We assessed severity independently and then discussed any differences in score to agree on a rating.

4.1 Differential Access

The most basic factor that determined who would and would not be able to access the Google Relate application was, as expected, the availability of a smartphone that would match the minimum specification outlined on Google Play. Although all 20 participants had access to a mobile phone before they signed up to participate in the study, one participant did not own the phone and could only access a phone through her mother, and the other seven did not have a smartphone with the required specifications. Four had a feature phone, and the remaining three had smartphones with insufficient RAM to support the application. We provided adequate smartphones to participants as part of the study. Still, many mentioned that they would have struggled to afford one, especially as purchasing a new phone was not considered a priority, considering their current one was functional. None knew they could use a higher phone specification to support an application to help communication in daily life. Even when in possession of a mobile phone with the required characteristics, several participants who had older or cheaper phones, including P4, P5, P6, P14, and P16, faced issues at different stages as the application was more prone to slow starts, unexpected crashes and failures to download updated versions of the voice model.

Table 1: Frequency of Special Characters

Participant	Aetiology	Role	Severity	Phrases	Custom Cards	Relate WER
P1	Cerebral Palsy	University Student	Moderate	679	382	26.2
P2	Cerebral Palsy	Own Business	Mild	520	20	20.8
P3	Cerebral Palsy	University Student	Moderate	1301	10	19.1
P4	Cerebral Palsy	Professional	Mild	500	35	14.2
P5	Cleft Palate	University Student	Mild	835	17	19.4
P6	Cerebral Palsy	University Student	Mild	1003	17	9.6
P7	Cleft Palate	Professional	Moderate	500	32	14.0
P8	Laryngectomy	Retired	Severe	150	NA	NA
P9	Laryngectomy	Retired	Moderate	1200	0	20.6
P10	Parkinson's Disease	Professional	Moderate	1024	8	14.8
P11	Stammer/Stutter	University Student	Moderate	577	61	12.3
P12	Laryngectomy	Professional	Moderate	603	12	20.1
P13	Cerebral Palsy	Living at home	Profound	603	126	83.6
P14	Cleft Palate	Professional	Moderate	500	4	16.8
P15	Laryngectomy	Professional	Moderate	504	4	13.9
P16	Cerebral Palsy	Professional	Moderate	3125	8	21.3
P17	Cerebral Palsy	University Student	Mild	512	16	11.3
P18	Cerebral Palsy	Professional	Severe	5125	16	17.4
P19	Stroke with Aphasia	Retired	Moderate	500	0	35.2
P20	Stroke with Aphasia	Retired	Severe	261	NA	NA

Another factor which affected the extent to which participants were able to access Google Relate was internet connectivity. Although the rate of mobile internet penetration in Ghana is rapidly increasing, the reliability of internet connection is generally poor, especially outside Accra, where P1, P13, P14, P16, P17, P18 live. Availability and cost of internet connection affected overall application access and specific features. P1, P11, P17, P18 reported concerns over the amount of data consumed by Google Relate. In particular, in one of our messages, P1 mentioned:

“Relate has a high data consumption in my experience. Also it needs very efficient data for functioning which is sometimes an issue in the northern part where I live”. - P1

To navigate this issue P6 devised a clever strategy that, with her permission, we have since shared with other participants. Although the Repeat and Assistant function require an internet connection the Listen function does not. In one of the videos shared on WhatsApp she explains:

“I can use Listen and speak normally so Relate writes it. When I finish I press stop and then the voice image (the speaker icon at the top left of the screen) and it reads it out to people. It is the same (i.e. the same function as Repeat) but it is free. Is also useful that it is written. Once I was in a shop and I wanted to ask something to the woman there. She could not really understand the English because it was fast and the shop was loud, but she could see the words on the screen and understand enough to serve me” - P6

The example above also shows another key aspect determining the level of access participants had to Google Relate: language. As mentioned previously the application is only available in English, meaning that the user has to be able to speak and understand English. However, this limitation does not just apply to the user

themselves, but anyone they wish to communicate with leveraging Relate. P11, P15, P16 and P19 all mentioned how in many situations they could not use Relate effectively as their conversational partners did not speak English. Others such as P1, P3, P5, P14 pointed to frequent mistakes made by the app when using local names or terms common in Ghana but not featured in the existing language model.

“Because of the location of my shop most of the customers communicate in the vernacular language which is a challenge because the app doesn't recognise my speech when I speak in the local dialect.” - P19

Finally, access to Google Relate is also mediated by the severity of the user's dysarthria or dysphonia and the presence of other conditions which might impair their ability to complete the number of recordings necessary for the creation of the bespoke voice model, as well as the degree of accuracy the voice model would have once created. Both P13 and P20 represent poignant examples of this as they faced, and still face, significant difficulties in recording the required number of sentences to be able to access Relate's various functions. P20's aphasia makes it challenging for him to read or repeat unfamiliar sentences, causing frustration and fatigue, meaning that only a small number of sentences can be recorded during his sessions with the SLT. Similarly, P13's profound dysarthria leads to difficulties pronouncing long sentences and unfamiliar words present in many of the “standard phrases” used in the Record tab to collect the speech samples needed for the creation of the model. To navigate this challenge P13 and her mother created a large number of custom cards which better aligned with her communication needs and capabilities.

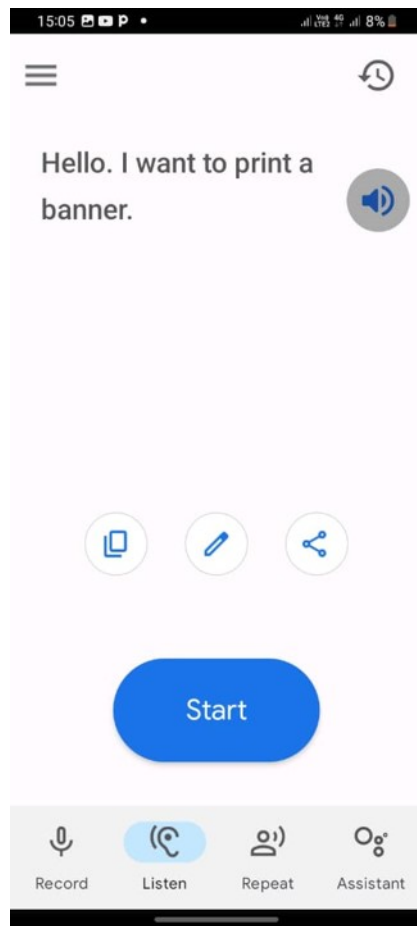


Figure 2: Screenshot of the Listen screen used by P7

4.2 Differential Capacity

The process of training Google Relate in building the customised voice model required participants to read out loud at least 500 English sentences, which implies a certain degree of literacy. As it was an inclusion criteria for the study, all our participants were able to understand and communicate in English. However, English was not the primary language for communication for several of them, and some including P10, P12, P13, P18, P19, P20 faced difficulties reading, especially when sentences included unfamiliar words (e.g. nachos or sushi), had meanings for which they lacked the context (e.g. *In junior high we went on a trip to Washington DC*), or simply did not make much sense (e.g. *Fish are quiet, they don't say anything*). The Record function offers the possibility to the user to have the sentence read out loud by the phone and repeat it, rather than having to read it from the text. However, we noticed several participants struggling with this option as the synthesised voice used in the application had an American accent, which was not necessarily easy to understand for participants, and used an American pronunciation which could differ from the one used by participants in everyday life.

P7: I cannot pronounce the world thermostat like it wants me

Researcher: Can you pronounce it for me?

P7: Thermostat. You see, it's wrong

Researcher: There is nothing wrong with it. You should not change how you speak for the app. If you try to read it like the voice plays it today, but then it's not how you say it next time it will not recognize it. If you pronounce the thermostat like this it's better for the app to know it. The phone should learn to understand you, not the other way around.

The exchange above shows how when there was a discrepancy between the pronunciation of the phone and their own, they were likely to incorrectly blame themselves. However, as pointed out by the researcher who carried out the training, the purpose of the application is to create a customised speech model based on the participant's regular speech pattern. When a person tries to change the way they speak to "please the app" they are more likely to introduce errors in the voice model created for them.

Understanding how the creation of the speech model works requires knowledge of how the application operates, which in turn is dependent on the person's digital literacy. The degree of digital literacy and contextual understanding of how Google Relate works also influenced the relative importance that people attributed to the creation of specific custom cards. Despite the time and effort involved in it, P1, P4, P7, P11 and P13 reported that they engaged in the creation of a larger number of custom cards because they understood how it could help them create a voice model that would be more contextually appropriate to their everyday communication ensuring that the application would correctly interpret the names of their family members, friends, important places, favourite foods, and commonly used words at work or school. In turn this could make Google Relate more useful in their daily life, which may lead to increased use and increased benefits.

Another key aspect that shaped the extent to which participants could use and benefit from the application was the degree of support they were able to leverage, particularly in earlier stages of learning or in case of technical difficulties. SLTs, friends and family members proved to be much more effective than the synthesised voice featured in the application in reading sentences out loud so participants could repeat them and complete the necessary 500 recordings (See Figure). P13's mother created a large number of custom cards that would be easier for her daughter to pronounce. P20's SLT also recently began to craft custom cards which include sentences from prayers that her client recites everyday as they are easier for him to remember despite his aphasia. Participants also mentioned how, the capacity to create custom cards, which is currently and individual effort, could be harnessed as a collective to improve the cultural relevance of the application. P1, P4 and P13 all stated that they would value an option that give them consent to share custom cards phrases with Google if this allowed other Ghanaian users to benefit from it, reducing the burden of custom card creation and increasing the recognition of important words and sentences.

Finally, one's capacity to extract benefit from the application in everyday life is hugely influenced by the opportunity to engage in communication with others in the first place. Google Relate, and other ASR for non-standard speech, are primarily intended as tools to aid listeners who are not familiar with the user to better understand them. But what happens when the stigma surrounding



Figure 3: P19 being helped by his SLT to record the sentences

their speech difficulties prevents people from engaging with others beyond their immediate social circles? P9 for example reported high satisfaction with Relate and stated that it can understand her well when she speaks. However, so far she has only been able to use the application with members of her family as she has limited contact with strangers as there is significant stigma surrounding people with Tracheo oesophageal puncture valve and she is more likely to rely on others to speak on her behalf when outside.

4.3 Differential Motivation

Our participants in the study, had various degrees of speech difficulty, lived in different parts of Ghana, came from a variety of educational, professional and social backgrounds, as well as having different life experiences. Their specific characteristics, personalities, and contextual situations determined their desire to use Relate for a variety of purposes. Many who worked or owned small businesses like P14 were especially happy that they could use Relate when working on a stall at the local market as *“the customers easily understood me when I use the app, and they can see the words if they cannot hear me”*. Others like P17 were delighted that they could use the keyboard function in university *“to write my essays when I am tired because writing is difficult as I cannot move my hand well, but*

now I can just say what I need and it writes it for me”. Those who had limited social interactions beyond their family because of stigma like P13 found joy in being able to *“talk with my best friend, my cousin [name] when he came for a visit”*.

What counts as success when using Relate appears to be a personal assessment, shaped by the lived experience of communication, whether it took place in public or private, involved complex interactions or simple exchanges, occurred with strangers or family members, the words used, and indeed whether it took place at all.

Like any other kind of new assistive technology that one might choose to try in a particular situation, participants understood that using Google Relate would involve a certain degree of risk taking, as they could not be sure of a successful outcome. The degree of risk that they were willing to take was influenced by a variety of factors including their own personalities, the perceived stakes of a particular situation and the alternatives available to them. A striking example of this is P4 a practising lawyer who had previously stopped attending her clients in court as judges had refused to hear her arguments labelling her as *“too difficult to understand”*. P4’s goal is to convince the judicial system to let her use Relate to argue her cases in a court of law, showing that with the right assistive technology, a judge can be able to understand her, regardless of

her speech difficulties. It should be stated that P is fully qualified to do this role. P4 is also an advocate for the rights of people with disabilities in the country and, at the time of writing, she is currently working with members of our team to ensure that she can use Google Relate to deliver a speech at an international conference in the coming months. Both these use-cases have high stakes, as failures of Relate could result in compromising the delivery of her speech in front of hundreds of people, or jeopardising her professional credibility in front of a judge. Discussions we had with P4 reveal how she is perfectly aware of these risks, but also of the potential benefits for herself and others, and of the lack of alternatives available to her.

“It’s a big job, you know, speaking in front of so many people. And I really want this app to carry out my legacy. If I can talk with it and show to all the other people with my condition that they can do this, they can achieve what they want.... My mission is that none of them has to struggle like me. I don’t want anyone to go through what I have. But until now I could not get my message across like I want, and I have tried many things...” P4

Interactions mediated by Relate were, of course, sometimes unsuccessful and the extent to which people would be comfortable to use it in everyday life varied on the basis of perceived benefit and other considerations. P12 mentioned that the app would misunderstand him from time to time, but less frequently than strangers, meaning that it could still deliver an advantage compared to relying on speech alone. P11, who had recently been diagnosed with Parkinson’s disease, found the app useful to use when he communicated with shopkeepers or other unfamiliar people, but it was hesitant to use it in work as not everyone was aware of his condition and he did not want to be outed as someone who needed technology to supplement his communication.

As communication is, by its nature, a collaborative act that involves the individual as well as the person they are communicating with, it was not just the motivation of the user which affected the use of Relate, but also the motivation of their conversational partner. As highlighted before, Google Relate is not an application that helps the user to speak, but one that helps the listener to understand. It is important that both speaker and listener are provided an opportunity to understand what Relate is and how it can be used to support conversation. When the listener did not understand what Relate was, appeared unsure how to adapt their conversation style to include it, or not willing to accept Relate in the conversation, social interactions could derail - focusing on Relate as opposed to the subject at hand, or fail.

We witnessed one of these failures during one of the ethnographic observations conducted with P5 who worked as a mobile money operator to pay for her nursing studies at the local university. The sequence is summarised below:

As the client approaches P5 attempts to use Relate’s Repeat function to facilitate communication and asks: *“Do you want to deposit or withdraw money?”*. The customer does not wait for Relate to repeat the sentence out loud and misunderstands P5 thinking that she asked *“Do you want to withdraw?”*. She responds: *“No, No I am not withdrawing I just want to charge!”* P5 tries to use Relate again and asks *“How much do you want to charge?”*. The customer does not wait for Relate to repeat and seems to misunderstand again. She shows P5 some notes she wants to deposit on her account and

says *“You see I want to charge my money”*. P5 realises the interaction with Relate has failed and she puts away her phone and shows the mobile money terminal to the customer, which displays the charging interface, so that the customer can see that she had understood from the beginning. The transaction is then completed without further mistakes.

The vignette above shows how the lack of understanding from the client of how Google Relates operate and how the conversation should be changed to make space for it hinders the exchange and the situation quickly unravels. It rests on P5’s shoulders to quickly adapt to the situation and find another way to help the client understand her so that they can complete the transaction. However if the client had an opportunity to understand how Relate works and how it can support communication the outcome may have been different.

One final example shows the potential for Relate to support advocacy by just being present in the room. P1 is the first person with cerebral palsy that has been accepted onto her course at the university she attends in the north of Ghana. She met with her head of department and used Relate Repeat in a conversation. The head of department commented that he wanted to support her to use Relate in class and to encourage other prospective students with similar communication differences to apply to university. He wants his department to be seen as inclusive and act as an example to others. This conversation happened because P1 had the courage to show how she can be better understood using technology, and in turn provided the head of department an opportunity to consider the practicalities of inclusiveness.

5 DISCUSSION

Our work represents, to the best of our knowledge, the first example of research examining the use of a free mobile-based ASR application for people with non standard speech in Ghana with the goal of enhancing communication equity. Our findings highlight the amplification mechanisms that determine who does and does not benefit from the use of the application revolve around (i) Differential Access: including ownership of a suitable smartphone with appropriate specification, internet connectivity, and language spoken; (ii) Differential capacity: including literacy, availability of support, understanding of how the technology works, and ability to interact with strangers; (iii) Differential motivation: including life experiences, willingness to risk, perceived benefits, and attitudes of conversational partners. Based on our results we highlight three key design implications for future technological systems aiming to promote communication equity in Ghana and other regions of the Global South for people who have difficulty articulating speech.

5.1 Adapting language to the context

Although English is considered the official language and it is reportedly the most commonly spoken, Ghana is a multilingual country with over 80 different languages being spoken [26]. Largely as a result of colonialist practices that have eradicated the use of indigenous languages, and the people who spoke them, the situation is similar in many other African countries where English is proposed as a national lingua franca that unifies the bureaucracy of the country [69]. Regardless of these colonialist and neocolonialist practices there are millions of people in Ghana and in other countries in

Africa who do not speak English. Technological systems aiming to promote communication equity in countries with such a diverse palette of languages cannot fail to consider these aspects, especially considering that people who speak indigenous languages are likely to belong to the most marginalised groups in society [6]. Our participants who lived and worked in poorer areas were the ones who most struggled to use Relate in daily conversations, as they lacked communication partners who could understand English.

Furthermore, even those who used English as their primary language in everyday communication found that Google Relate was prone to make mistakes when they used local words and expressions that did not exist in American English, the language on which the ASR model was based. Ghanaian English has a distinct pronunciation and a broad local vocabulary that has developed over decades of use and contextualization [37]. On the one hand it could be argued that the creation of a bespoke speech model should address issues of pronunciation. However, we observed that during the recording of their own voice samples, when participants heard the application enunciate words with an unfamiliar, and contextually incorrect, pronunciation, they were likely to attempt to emulate it thinking that the application was speaking in “proper English”. This is not only demeaning for the person, but also counterproductive, as the speech recognition model would effectively be built on samples that do not match the person’s actual speech in everyday life.

When it came to building a specific vocabulary to ensure that Google Relate could understand frequently used local names and keywords, the only viable strategy available to users at the moment is the creation of Custom cards through the dedicated function. Participants who created a greater number of custom cards were most likely to report greater satisfaction with the application, even when their WER were still relatively high. This is because the application was able to recognise the words that were important to them. However, the process of custom card creation is cumbersome, often resulting in need for support, and requires participants to understand its relative importance for the creation of a contextualised language model. The integration of new generative AI based on large language models could at least in part improve this aspect by enabling participants to automatically create custom cards based on a set of chosen words and ensuring the correct configuration featuring the word at the start, middle, and end to maximise recognition. Generative models could also create easier sentences to record for people by simplifying or shortening the sentences for participants with more profound communication difficulties, as well as creating contextualised sentences that match the vocabulary of participants similarly to the the work done by the mother of P13 or the SLT of P20 [89]. Beyond generative models it might also be possible to create designs which are more adaptable to the wide range of English that is spoken globally by incorporating this as a design requirement from the outset. Knowing English will vary in how it is spoken, and acknowledging the privileged position of having large data sets of only a subset of the range of English we can design in adaptability from the outset. In this case perhaps custom cards would have been made easier to use or indeed the word base used to initially train the application could have been edited. More widely we can look to create better datasets for training of languages and versions of languages.

5.2 Not just about the user

As we mentioned several times throughout the paper, Google Relate as well as other ASRs are not technologies that are meant to change the communication strategies of people with dysarthria or dysphonia, but are tools meant to support their communication partners to understand them better. In particular these systems are supposed to be used in conversations with strangers and unfamiliar partners, as people close to the users are likely to have learned to understand how the persons speak and do not necessarily benefit from ASR [9, 35]. As a result, if the goal is the successful introduction of these technologies we cannot disregard the importance of improving awareness and training at a more systemic level. Traditionally assistive technology is rolled out with training, however, this is focussed on the user which is normally the person with the speech impairment. In Google Relate’s case we are helping people with hearing and understanding. This changes the nature of the traditional assistive technology training paradigm. As seen in some of the vignettes presented the use of ASR can change the dynamic of the conversation including adding pauses and altering the flow of an exchange. If the other person involved in the conversation does not understand or refuses to accept these “new rules” the interaction is bound to fail. On the other hand, like the head of the department in P1’s department, if the conversational partner can be convinced to listen, it is possible that a successful interaction might lead to broader reconsiderations of inclusion and accessibility at systemic level. During our conversation with P1 and her SLT, we have hypothesized that the interaction between her and director of the department was, at least in part, made possible by the perceived power differential created by a researcher from a Global North institution accompanying one of the students. Although there is no possibility for us to be sure if this is the case, the assumption is based on on the contextual knowledge of the participant and local SLT, who have deep personal experiences of the barriers and discrimination faced by people with communication difficulties. On the surface, the ability of certain members of our team to leverage the power differential to support a participant felt like a positive opportunity as P1’s activism and advocacy could substantially improve the accessibility of student experience for herself and others. At the same time, it highlights a deep injustice in a system which refuses to listen to her in the first place, requiring the presence of a white researcher in the room who had no contextual knowledge of how to deliver systemic change in the university, but simply acted as an enabler to the meeting in virtue of its perceived social status.

Systemic change is sorely needed for people with communication difficulties in Ghana, many of our participants had been victims of severe stigmatisation, abuse, and marginalisation for most of their life [8, 63]. Some of them were limited in their abilities to use Relate because of the social stigma that limited their interactions beyond their immediate circles of family and friends. Technology such as cannot create resources or good intent where they do not exist, and this has to be recognised to make space for broader interventions that can create the right conditions for technology to amplify efforts towards communication equity. What HCI researchers and developers of technologies such as Relate can do instead is attempt

to amplify the efforts of individuals like P1 and P4 who are doing the actual work of changing wills, hearts and minds [88].

Finally, it is essential to notice how a key enabler or participants success was access to support being in the form of family members that could help to create custom cards, SLTs and researchers to help with set up and troubleshooting, and participants themselves who shared strategies to bypass technological limitations. Building capacity amongst providers such as SLTs, teachers, community leaders and advocates is the only way to promote scalability of these technologies, which in turn can contribute to drive social acceptance and lead to greater and more equitable benefits amongst potential users [47, 83].

5.3 A tool not a solution

What we have found is that Project Relate is useful in certain contexts, however it is not yet a universal solution to communication difficulties between people when one or more have a speech impairment. This is because there are challenges in language, data consumption, functioning speed, contextual understanding, and ease of use were to be addressed. From an HCI perspective these aspects are important considerations within the design process. Through their everyday practices people with non-standard speech have developed personalised communication strategies that do not involve the use of ASR and, depending on the contextual situation, they might still prefer to more conventional methods that do not require the use of a technological intermediary [82, 83]. Therefore future work needs to recognise how people with non-standard speech currently communicate, and then how this can be enhanced by technology. It is not necessarily the case that technology needs to take the place of the current communication pattern, but could instead be a helpful adjunct whenever chosen by the person communicating. Furthermore, as a community we should be careful to not unintentionally change the speech patterns of people, to normalise them through word choice or sentence structure or even articulation to fit a normative model which is not their local culture. Finally, we noted the need for training and that the challenges posed above resulted even with considerable training and support. Given applications such as Google Relate can be downloaded by anyone, they have great potential to widen the availability of assistive technology. However as demonstrated in this research, this will only be possible with a high standard of training and support, which will most likely involve some level of clinical input, and this in turn will require knowledge that the new application exists.

We invite HCI and accessibility researchers as well as designers and developers to consider these aspects and engage in more transparent reflections on the limitations of current ASRs and AACs, implications that the use of these technologies has for people's everyday communication. After all, when barriers in social interactions involving speech occur, people with communication difficulties and their conversational partners will find solutions to them, sometimes these solutions will involve the use of technology as a tool, but other times they will not.

6 CONCLUSION

In conclusion, our research into the use of free mobile-based Automatic Speech Recognition (ASR), Project Relate applications, for

individuals with non-standard speech in Ghana has illuminated a multifaceted landscape of challenges and opportunities. To promote communication equity effectively, it is imperative to address differential access, recognizing the need for equitable access to technology and accounting for linguistic diversity. Additionally, the richness of local languages must be acknowledged, and ASR models should be adapted to understand and respect these linguistic nuances. Future studies should seek to unpack the impact of linguistic, cultural, and other social and environmental factors on ASR technologies across a variety of geographies in both the Global South and Global North to build a more nuanced perspective and highlight opportunities for reciprocal learning, rather than relying on colonial expectations of technology transfer. Finally, our study emphasises the significance of societal change beyond technology adoption. ASR tools such as Project Relate should not be seen as stand-alone solutions but as instruments that can reshape communication dynamics. Raising awareness and providing training at a systemic level is crucial for successful integration into various social contexts, acknowledging the transformative potential of these technologies.

ACKNOWLEDGMENTS

This work was funded by the Assistive Technology 2030 project, funded by the United Kingdom Foreign Commonwealth Development Office (FCDO; UK Aid) and administered by the Global Disability Innovation Hub. We sincerely want to thank all the participants and speech and language therapists that shared and continue to share with us their experience, skills, and wisdom. Without their outstanding contribution, this research would simply not exist.

REFERENCES

- [1] [n. d.]. Google for Africa 2022. <https://nextbillionusers.withgoogle.com/events/google-4-africa-2022>
- [2] [n. d.]. Project Relate – Google Research. <https://sites.research.google/relate/>
- [3] 2023. Project Euphonia. <https://sites.research.google/euphonia/publications/>
- [4] Muhammad Sadi Adamu. 2023. No more “solutionism” or “saviourism” in futuring African HCI: A manifesto. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–42.
- [5] Blessing F. Adeoye. 2022. The Deaf Community in Ghana as Technology Adopters. (May 2022). <https://openoregon.pressbooks.pub/ghanaiandeaf/chapter/chapter-2/> Book Title: Signed languages, interpreting, and the Deaf Community in Ghana and West Africa.
- [6] Grace Ampomaa Afrifa, Jemima Asabea Anderson, and Gladys Nyarko Ansah. 2019. The choice of English as a home language in urban Ghana. *Current Issues in Language Planning* 20, 4 (2019), 418–434.
- [7] Paul Kwaku Larbi Anderson, Johannes Schädler, and Lars Wissenbach. [n. d.]. Information and Communication Technology in Rural Healthcare and Social Welfare Service Provision in Ghana-Prospects in the Face of Social Inequalities. *social welfare* 1 ([n. d.]), 2.
- [8] Michael Baffoe. 2013. Stigma, discrimination & marginalization: Gateways to oppression of persons with disabilities in Ghana, West Africa. *Journal of Educational and Social Research* 3, 1 (2013), 187–198.
- [9] V Balaji and G Sadashivappa. 2015. Speech disabilities in adults and the suitable speech recognition software tools-a review. In *2015 International Conference on Computing and Network Communications (CoCoNet)*. IEEE, 559–564.
- [10] Laura J Ball, David R Beukelman, and Gary L Pattee. 2004. Communication effectiveness of individuals with amyotrophic lateral sclerosis. *Journal of Communication Disorders* 37, 3 (2004), 197–215.
- [11] Fabio Ballati, Fulvio Corno, and Luigi De Russis. 2018. “Hey Siri, do you understand me?”: Virtual Assistants and Dysarthria. In *Intelligent Environments 2018*. IOS Press, 557–566.
- [12] Giulia Barbareschi, Catherine Holloway, Katherine Arnold, Grace Magomere, Wycliffe Ambeyi Wetende, Gabriel Ngare, and Joyce Olenja. 2020. The social network: How people with visual impairment use mobile phones in kibera, Kenya. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

- [13] Giulia Barbareschi, Ben Oldfrey, Long Xin, Grace Nyachomba Magomere, Wycliffe Ambeyi Wetende, Carol Wanjira, Joyce Olenja, Victoria Austin, and Catherine Holloway. 2020. Bridging the Divide: Exploring the use of digital and physical technology to aid mobility impaired people living in an informal settlement. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–13.
- [14] David R Beukelman, Susan Fager, Laura Ball, and Aimee Dietz. 2007. AAC for adults with acquired neurological conditions: A review. *Augmentative and alternative communication* 23, 3 (2007), 230–242.
- [15] David R Beukelman, Pat Mirenda, et al. 1998. *Augmentative and alternative communication*. Paul H. Brookes Baltimore.
- [16] F Biadys, RJ Weiss, PJ Moreno, D Kanvesky, and Y Jia. [n.d.]. Parrottron: An end-to-end speech-to-speech conversion model and its applications to hearing-impaired speech and speech separation. arXiv 2019. *arXiv preprint arXiv:1904.04169* (n.d.).
- [17] Stephanie A Borrie and Kaitlin L Lansford. 2021. A perceptual learning approach for dysarthria remediation: An updated review. *Journal of Speech, Language, and Hearing Research* 64, 8 (2021), 3060–3073.
- [18] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.
- [19] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 18, 3 (2021), 328–352.
- [20] Mary Bucholtz and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse studies* 7, 4–5 (2005), 585–614.
- [21] Irene Calvo, Peppino Tropea, Mauro Viganò, Maria Scialla, Agnieszka B Cavalcante, Monika Grajzer, Marco Gilardone, and Massimo Corbo. 2021. Evaluation of an automatic speech recognition platform for dysarthric speech. *Folia Phoniatrica et Logopaedica* 73, 5 (2021), 432–441.
- [22] Richard Cave and Steven Bloch. 2021. The use of speech recognition technology by people living with amyotrophic lateral sclerosis: A scoping review. *Disability and Rehabilitation: Assistive Technology* (2021), 1–13.
- [23] Zoë Charlotte Clarke, Simon Judge, Kate Fryer, Stuart Cunningham, Jonathan Toogood, and Mark S Hawley. 2023. A qualitative study exploring the effect of communicating with partially intelligible speech. *Augmentative and Alternative Communication* (2023), 1–13.
- [24] Catherine Crowley, Miriam Baigorri, Clement Ntim, Belinda Bukari, Albert Oseiabgyina, Emmanuel Kitcher, Albert Paintsil, Opoku Ware Ampomah, and Anthony Laing. 2013. Collaborations to address barriers for people with communication disabilities in Ghana: Considering the World Report on Disability. *International Journal of Speech-Language Pathology* 15, 1 (2013), 53–57.
- [25] Shakila Dada, Yvonne Murphy, and Kerstin Tönsing. 2017. Augmentative and alternative communication practices: A descriptive study of the perceptions of South African speech-language therapists. *Augmentative and Alternative Communication* 33, 4 (2017), 189–200.
- [26] Mary E Kropp Dakubu. 2015. *The languages of Ghana*. Routledge.
- [27] Luigi De Russis and Fulvio Corno. 2019. On the impact of dysarthric speech on contemporary ASR cloud platforms. *Journal of Reliable Intelligent Environments* 5 (2019), 163–172.
- [28] Cornelius Debuur, Paul Welaga, George Wak, and Abraham Hodgson. 2010. Self-reported health and functional limitations among older people in the Kassena-Nankana District, Ghana. *Global health action* 3, 1 (2010), 2151.
- [29] Benoît Favre, Kyla Cheung, Siavash Kazemian, Adam Lee, Yang Liu, Cosmin Munteanu, Ani Nenkova, Dennis Ochei, Gerald Penn, Stephen Tratz, et al. 2013. Automatic human utility evaluation of ASR systems: Does WER really predict performance?. In *INTERSPEECH*. 3463–3467.
- [30] James Gee. 2015. *Social linguistics and literacies: Ideology in discourses*. Routledge.
- [31] Kamala Gollakota, James B Pick, and Manju Singh. 2022. Does purpose of use matter? Influences on developmental use versus personal use by low-income farmers. *Information Technology for Development* 28, 1 (2022), 111–136.
- [32] Jordan R Green, Robert L MacDonald, Pan-Pan Jiang, Julie Cattiau, Rus Heywood, Richard Cave, Katie Seaver, Marilyn A Ladewig, Jimmy Tobin, Michael P Brenner, et al. 2021. Automatic Speech Recognition of Disordered Speech: Personalized Models Outperforming Human Listeners on Short Phrases.. In *Interspeech*. 4778–4782.
- [33] Foad Hamidi, Patrick Mbullo Owuor, Michaela Hynie, and Melanie Baljko. 2022. "Knowledge Comes Through Participation": Understanding Disability through the Lens of DIY Assistive Technology in Western Kenya. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–25.
- [34] Sally Hartley. 1998. Service development to meet the needs of people with communication disabilities in developing countries. *Disability and rehabilitation* 20, 8 (1998), 277–284.
- [35] Mark S Hawley, Stuart P Cunningham, Phil D Green, Pam Enderby, Rebecca Palmer, Siddharth Sehgal, and Peter O'Neill. 2012. A voice-input voice-output communication aid for people with severe speech impairment. *IEEE Transactions on neural systems and rehabilitation engineering* 21, 1 (2012), 23–31.
- [36] Catherine Holloway and Giulia Barbareschi. 2022. Exploring Geographies: DIX in the Global South. In *Disability Interactions: Creating Inclusive Innovations*. Springer, 71–90.
- [37] Magnus Huber. 2008. Ghanaian English: Phonology. *Varieties of English* 4 (2008), 67–92.
- [38] Lorraine Hudson, Clement Amponsah, Josephine Ohenewa Bampoe, Julie Marshall, Nana Akua Victoria Owusu, Khalid Hussein, Jess Linington, Zoe Banks Gross, Jane Stokes, and Róisín McNaney. 2020. Co-designing digital tools to enhance speech and language therapy training in Ghana. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [39] Caroline Jagoe, Caitlin McDonald, Minerva Rivas, and Nora Groce. 2021. Direct participation of people with communication disabilities in research on poverty and disabilities in low and middle income countries: A critical review. *Plos one* 16, 10 (2021), e0258575.
- [40] Ensa Johnson, Stefan Nilsson, and Margareta Adolfsson. 2015. Eina! Ouch! Eish! Professionals' perceptions of how children with cerebral palsy communicate about pain in South African school settings: Implications for the use of AAC. *Augmentative and Alternative Communication* 31, 4 (2015), 325–335.
- [41] Jasmine Jones and Joyjeet Pal. 2015. Counteracting dampeners: understanding technology-amplified capabilities of people with disabilities in Sierra Leone. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. 1–10.
- [42] Alexander B Kain, John-Paul Hosom, Xiaochuan Niu, Jan PH Van Santen, Melanie Fried-Oken, and Janice Staehely. 2007. Improving the intelligibility of dysarthric speech. *Speech communication* 49, 9 (2007), 743–759.
- [43] Harsha Kathard and Mershen Pillay. 2013. Promoting change through political consciousness: A South African speech-language pathology response to the World Report on Disability. *International journal of speech-language pathology* 15, 1 (2013), 84–89.
- [44] Susan Koch Fager, Melanie Fried-Oken, Tom Jakobs, and David R Beukelman. 2019. New and emerging access technologies for adults with complex communication needs and severe motor impairments: State of the science. *Augmentative and Alternative Communication* 35, 1 (2019), 13–25.
- [45] Allison Koencke, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R Rickford, Dan Jurafsky, and Sharad Goel. 2020. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences* 117, 14 (2020), 7684–7689.
- [46] Per Ola Kristensson, James Lilley, Rolf Black, and Annalu Waller. 2020. A design engineering approach for quantitatively exploring context-aware sentence retrieval for nonspeaking individuals with motor disabilities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [47] Natasha Layton, Caitlin Murphy, and Diane Bell. 2018. From individual innovation to global impact: the Global Cooperation on Assistive Technology (GATE) innovation snapshot as a method for sharing and scaling. *Disability and Rehabilitation: Assistive Technology* 13, 5 (2018), 486–491.
- [48] Jimin Lee, Katherine C Hustad, and Gary Weismer. 2014. Predicting speech intelligibility with a multiple speech subsystems approach in children with cerebral palsy. *Journal of Speech, Language, and Hearing Research* 57, 5 (2014), 1666–1678.
- [49] Shansong Liu, Mengzhe Geng, Shoukang Hu, Xurong Xie, Mingyu Cui, Jianwei Yu, Xunying Liu, and Helen Meng. 2021. Recent progress in the CUHK dysarthric speech recognition system. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 29 (2021), 2267–2281.
- [50] Rena Lyons and Sue Roulstone. 2017. Labels, identity and narratives in children with primary speech and language impairments. *International Journal of Speech-Language Pathology* 19, 5 (2017), 503–518.
- [51] Bob MacDonald, Pan-Pan Jiang, Julie Cattiau, Rus Heywood, Richard Cave, Katie Seaver, Marilyn Ladewig, Jimmy Tobin, Michael Brenner, Philip Q Nelson, et al. 2021. Disordered speech data collection: lessons learned at 1 million utterances from project euphonia. (2021).
- [52] Beza Merid, Maria Cielito Robles, Brahmajee K Nallamothu, Mark W Newman, and Lesli E Skolarus. 2022. Viewing Mobile Health Technology Design Through the Lens of Amplification Theory. *JMIR mHealth and uHealth* 10, 6 (2022), e31069.
- [53] Jiří Mertl, Eva Žáčková, and Barbora Řepová. 2018. Quality of life of patients after total laryngectomy: the struggle against stigmatization and social exclusion using speech synthesis. *Disability and Rehabilitation: Assistive Technology* 13, 4 (2018), 342–352.
- [54] Josh Meyer, Lindy Rauchenstein, Joshua D Eisenberg, and Nicholas Howell. 2020. Artie bias corpus: An open dataset for detecting demographic bias in speech applications. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*. 6462–6468.
- [55] Taniya Mishra, Andrej Ljolje, and Mazin Gilbert. 2011. Predicting human perceived accuracy of ASR systems. In *Twelfth Annual Conference of the International Speech Communication Association*.
- [56] Meredith Moore, Hemanth Venkateswara, and Sethuraman Panchanathan. 2018. Whistle-blowing ASRs: Evaluating the need for more inclusive speech recognition systems. *Interspeech* 2018 (2018).
- [57] Wakisa Mulwafu, Robbert Ensink, Hannah Kuper, and Johannes Fagan. 2017. Survey of ENT services in sub-Saharan Africa: little progress between 2009 and

2015. *Global health action* 10, 1 (2017), 1289736.
- [58] Monica Murero, Salvatore Vita, Andrea Mennitto, and Giuseppe D'Ancona. 2020. Artificial intelligence for severe speech impairment: innovative approaches to AAC and communication. In *PSYCHOBIT*.
- [59] Nimisha Muttiah, Jessica Gormley, and Kathryn DR Drager. 2022. A scoping review of augmentative and alternative communication (AAC) interventions in low-and middle-income countries (LMICs). *Augmentative and Alternative Communication* 38, 2 (2022), 123–134.
- [60] Anirudh Nagraj, Ravi Kuber, Foad Hamidi, and Raghavendra SG Prasad. 2021. Investigating the navigational habits of people who are blind in india. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–10.
- [61] Esther Nathanson. 2017. Native voice, self-concept and the moral case for personalized voice technology. *Disability and rehabilitation* 39, 1 (2017), 73–81.
- [62] Fayika Farhat Nova, MD Rashidujaman Rifat, Pratyasha Saha, Syed Ishtiaque Ahmed, and Shion Guha. 2019. Online sexual harassment over anonymous social media in Bangladesh. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*. 1–12.
- [63] Joseph Ocran. 2022. "There is something like a barrier": Disability stigma, structural discrimination and middle-class persons with disability in Ghana. *Cogent Social Sciences* 8, 1 (2022), 2084893.
- [64] Emel Oguz-Akarsu, Rifat Ozpar, Haci Mirzayev, Nilufer Aylin Acet-Ozturk, Bahattin Hakyemez, Dane Ediger, Necdet Karli, and Pandemic Study Team. 2020. Guillain-Barré syndrome in a patient with minimal symptoms of COVID-19 infection. *Muscle & nerve* 62, 3 (2020), E54–E57.
- [65] Joshua Annor Osam, Maxwell Peparh Opoku, Joslin Alexei Dogbe, William Nketsia, and Charles Hammond. 2021. The use of assistive technologies among children with disabilities: The perception of parents of children with disabilities in Ghana. *Disability and Rehabilitation: Assistive Technology* 16, 3 (2021), 301–308.
- [66] Joyojeet Pal, Tawfiq Ammari, Ramaswami Mahalingam, Ana Maria Huaita Alfaro, and Meera Lakshmanan. 2013. Marginality, aspiration and accessibility in ICTD. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1*. 68–78.
- [67] Joyojeet Pal, Anandhi Viswanathan, Priyank Chandra, Anisha Nazareth, Vaishnav Kameswaran, Hariharan Subramonyam, Aditya Johri, Mark S Ackerman, and Sile O'Modhrain. 2017. Agency in assistive technology adoption: visual impairment and smartphone use in Bangalore. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 5929–5940.
- [68] James Atta Peparh, Eric Atsu Avorkpo, and Evans Kulu. 2022. People with disability and access to financial services: Evidence from Ghana. *Regional Science Policy & Practice* (2022).
- [69] Kwesi Kwaa Prah. 2002. The burden of English in Africa: from colonialism to neo-colonialism. *TESL-EJ (Teaching English as a Second or Foreign Language–Electronic Journal)* 6, 1 (2002).
- [70] Kacper Radzikowski, Le Wang, Osamu Yoshie, and Robert Nowak. 2021. Accent modification for speech recognition of non-native speakers using neural style transfer. *EURASIP Journal on Audio, Speech, and Music Processing* 2021, 1 (2021), 1–10.
- [71] Ananditha Raghunath, Laurel Krovetz, Hosea Mpogole, Henry Mulisa, Brian Dillon, and Richard Anderson. 2023. From Grasshoppers to Secondhand Cars: Understanding the Smartphone-Enabled Marketplace in Peri-urban Tanzania. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [72] Kristin Rosen and Sasha Yampolsky. 2000. Automatic speech recognition and a review of its functioning with dysarthric speech. *Augmentative and Alternative Communication* 16, 1 (2000), 48–60.
- [73] Hannah P Rowe, Sarah E Gutz, Marc F Maffei, Katrin Tomanek, and Jordan R Green. 2022. Characterizing Dysarthria Diversity for Automatic Speech Recognition: A Tutorial From the Clinical Perspective. *Frontiers in Computer Science* 4 (2022), 770210.
- [74] Manika Saha, Delvin Varghese, Tom Bartindale, Shakuntala Haraksingh Thilsted, Syed Ishtiaque Ahmed, and Patrick Olivier. 2022. Towards Sustainable ICTD in Bangladesh: Understanding the Program and Policy Landscape and Its Implications for CSCW and HCI. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–31.
- [75] Theresa Schölderle, Anja Staiger, Renée Lampe, Katrin Strecker, and Wolfram Ziegler. 2016. Dysarthria in adults with cerebral palsy: Clinical presentation and impacts on communication. *Journal of Speech, Language, and Hearing Research* 59, 2 (2016), 216–229.
- [76] Suraj Singh Senjam, Souvik Manna, and Covadonga Bascaran. 2021. Smartphones-based assistive technology: accessibility features and apps for people with visual impairment, and its usage, challenges, and usability testing. *Clinical optometry* (2021), 311–322.
- [77] Araba Sey. 2011. 'We use it different, different': Making sense of trends in mobile phone use in Ghana. *New media & society* 13, 3 (2011), 375–390.
- [78] Araba Sey and Peppino Ortoleva. 2014. All work and no play? Judging the uses of mobile phones in developing countries. *Information Technologies & International Development* 10, 3 (2014), pp–1.
- [79] Barbara Shadden. 2005. Aphasia as identity theft: Theory and practice. *Aphasiology* 19, 3–5 (2005), 211–223.
- [80] Seyed Reza Shahamiri. 2021. Speech vision: An end-to-end deep learning-based dysarthric automatic speech recognition system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), 852–861.
- [81] Donghoon Shin, Jaeyoon Song, Seokwoo Song, Jisoo Park, Joonhwan Lee, and Soojin Jun. 2020. TalkingBoogie: collaborative mobile AAC system for non-verbal children with developmental disabilities and their caregivers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [82] Kristen Shinohara. 2012. A new approach for the design of assistive technologies: design for social acceptance. *ACM SIGACCESS Accessibility and Computing* 102 (2012), 45–48.
- [83] Kristen Shinohara and Jacob O Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 705–714.
- [84] Martine M Smith and Isobel Connolly. 2008. Roles of aided communication: Perspectives of adults who use AAC. *Disability and Rehabilitation: Assistive technology* 3, 5 (2008), 260–273.
- [85] Gabriel Synnaeve, Qiantong Xu, Jacob Kahn, Tatiana Likhomanenko, Edouard Grave, Vineel Pratap, Anuroop Sriram, Vitaliy Liptchinsky, and Ronan Collobert. 2019. End-to-end asr: from supervised to semi-supervised learning with modern architectures. *arXiv preprint arXiv:1911.08460* (2019).
- [86] Ryoichi Takashima, Tetsuya Takiguchi, and Yasuo Ariki. 2020. Two-step acoustic model adaptation for dysarthric speech recognition. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 6104–6108.
- [87] Kentaro Toyama. 2011. Technology as amplifier in international development. In *Proceedings of the 2011 iConference*. 75–82.
- [88] Kentaro Toyama. 2015. *Geek heresy: Rescuing social change from the cult of technology*. PublicAffairs.
- [89] Stephanie Valencia, Richard Cave, Krystal Kallarackal, Katie Seaver, Michael Terry, and Shaun K Kane. 2023. "The less I type, the better": How AI Language Models can Enhance or Impede Communication for AAC Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [90] Rama Adithya Varanasi, Divya Siddarth, Vivek Seshadri, Kalika Bali, and Aditya Vashistha. 2022. Feeling Proud, Feeling Embarrassed: Experiences of Low-income Women with Crowd Work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [91] Rajesh Veeraraghavan. 2013. Dealing with the digital panopticon: the use and subversion of ICT in an Indian bureaucracy. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1*. 248–255.
- [92] Annalu Waller. 2019. Telling tales: unlocking the potential of AAC technologies. *International journal of language & communication disorders* 54, 2 (2019), 159–169.
- [93] Margaret Walshe and Nick Miller. 2011. Living with acquired dysarthria: the speaker's perspective. *Disability and rehabilitation* 33, 3 (2011), 195–203.
- [94] Mary Wickenden. 2013. Widening the SLP lens: How can we improve the wellbeing of people with communication disabilities globally. *International Journal of Speech-Language Pathology* 15, 1 (2013), 14–20.
- [95] World Bank. 2018. Ghana Population. <https://www.worldbank.org/en/country/ghana/overview>
- [96] Susan Wyche, Nightingale Simiyu, and Martha E Othieno. 2016. Mobile phones as amplifiers of social inequality among rural Kenyan women. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 3 (2016), 1–19.
- [97] Susan Wyche, Charles Steinfield, Tian Cai, Nightingale Simiyu, and Martha E Othieno. 2016. Reflecting on video: Exploring the efficacy of video for teaching device literacy in rural Kenya. In *Proceedings of the eighth international conference on information and communication technologies and development*. 1–10.
- [98] Karen Wylie, Bronwyn Davidson, Julie Marshall, Josephine Ohenewa Bampoe, Clement Amponsah, and Lindy McAllister. 2020. Community service providers' roles in supporting communication disability rehabilitation in Majority World contexts: An example from Ghana. *International Journal of Speech-Language Pathology* 22, 4 (2020), 414–424.
- [99] Karen Wylie, Lindy McAllister, Bronwyn Davidson, and Julie Marshall. 2013. Changing practice: Implications of the World Report on Disability for responding to communication disability in under-served populations. *International journal of speech-language pathology* 15, 1 (2013), 1–13.
- [100] Karen Wylie, Lindy McAllister, Bronwyn Davidson, and Julie Marshall. 2018. Communication rehabilitation in sub-Saharan Africa: The role of speech and language therapists. *African Journal of Disability* 7, 1 (2018), 1–9.
- [101] Karen Wylie, Lindy McAllister, Bronwyn Davidson, Julie Marshall, Clement Amponsah, and Josephine Ohenewa Bampoe. 2017. Self-help and help-seeking for communication disability in Ghana: Implications for the development of communication disability rehabilitation services. *Globalization and health* 13 (2017), 1–13.

- [102] Junichi Yamagishi, Christophe Veaux, Simon King, and Steve Renals. 2012. Speech synthesis technologies for individuals with vocal disabilities: Voice banking and reconstruction. *Acoustical Science and Technology* 33, 1 (2012), 1–5.
- [103] Victoria Young and Alex Mihailidis. 2010. Difficulties in automatic speech recognition of dysarthric speakers and implications for speech-based applications used by the elderly: A literature review. *Assistive Technology* 22, 2 (2010), 99–112.