'I try to talk it white way': Australian Aboriginal English Speakers' Experiences with Voice Technologies

Anonymous ACL submission

Abstract

This article reports on a qualitative study investigating experiences of Indigenous people in Australia with voice technologies. Many Indigenous people in Australia speak Australian Aboriginal English, a variety which emerged as a result of forced contact between Indigenous communities and English speakers. It is widely used, yet is poorly supported by language technologies. Through Indigenous-led in-depth interviews, this research explores user experiences, attitudes, responses, and challenges encountered while interacting with various speech-enabled devices and applications using Australian Aboriginal English. We find that Indigenous users of ASR-powered technologies report poor experiences with the technologies, which they see as working better for non-Indigenous people. They report switching to speak like white people to get better ASR performance. Our findings reveal significant pressure to speak 'white way' when interacting with technology.

1 Introduction

Automated Speech Recognition (ASR) technologies have become ubiquitous in everyday life. These technologies are integrated into a wide range of commercial and freely available products, including virtual assistants, call centres, dictation products, and in-car interactive technologies (Aksënova et al., 2021). ASR technologies are rapidly becoming available in many languages, as well as in national varieties of international languages, accompanied by promises of inclusion for all the world (Javed et al., 2022). However there is evidence that ASR systems work poorly for minoritised language varieties, such as African American English (Koenecke et al., 2020) or South Asian English (Holt et al., 2024). Furthermore, speakers of minoritised language varieties experience negative psychological impacts when the technology fails, such as feeling that technology wasn't designed for them (Mengesha et al., 2021).

Of the 400+ ancestral Indigenous languages that existed in Australia before European colonisation (Bowern, 2023, p. 56), most are either extinct or critically endangered (Australian Government et al., 2020). This is due to practices of colonialism, which at various times included governments forcibly separating Aboriginal children from their families and forbidding Aboriginal children from speaking their languages at school (Tudor-Smith et al., 2024). As a result, most Indigenous people in Australia speak Australian Aboriginal English, making it a strong encoder of ethno-cultural identity (Rodríguez Louro and Collard, 2024). This indigenised variety of English integrates features from English and Aboriginal ancestral languages.

We adopt the position that Indigenous people in Australia should be supported to use Australian Aboriginal English—if they choose to—when using voice technologies. This premise raises an immediate question: what are the experiences of Australian Aboriginal English speakers using currently available voice technologies? This paper addresses this question, making the following contributions. First, we report on the first empirical study to explore the experiences of Indigenous people in Australia when using voice technologies. Second, we provide evidence that Australian Aboriginal English speakers report that ASR-powered technologies often fail for them when speaking their variety of English, causing feelings of frustration, and reproducing social and linguistic minoritisation. Third, we provide evidence that Indigenous people in Australia associate ASR performance with race, believing that voice technologies work better for white people, and modifying their voices to sound white to get better ASR performance. Throughout the project, we emphasise the importance of Indigenous participation and cultural safety.

2 Background and Related Work

2.1 Australian Aboriginal English

Australian Aboriginal English is a contact variety of English used by Indigenous people in Australia (Rodríguez Louro and Collard, 2021a). It is particularly prevalent across southern Australia where ancestral languages are less widely spoken. In these contexts, Australian Aboriginal English is used to strengthen Indigenous ways of being and communicating (Eades, 2012a). Australian Aboriginal English is the product of forced contact between Indigenous people and the English-speaking settler-colonialists. The English spoken by the British was a powerful instrument used to strip Indigenous communities of their ancestral rights (Malcolm, 2000, p. 125).

Because Australian Aboriginal English is English based, it is sometimes socially stigmatised, erroneously classified by laypeople as an incorrect version of English. As a result, L1 users of Australian Aboriginal English are often discriminated against in schooling and legal settings (Harkins, 1994; Eades, 2012b), perpetuating invisibility of the variety and leading to the false perception that its speakers are uneducated and rude (Eades, 2013). These negative attitudes are exacerbated because Australian Aboriginal English is primarily oral, with no nationally adopted guidelines as to how it should be written.

Indigenous people tend to find the use of direct questions awkward or offensive (Rodríguez Louro and Collard, 2024), and these pragmatic considerations continue in Australian Aboriginal English. Direct questions are also likely to elicit misleading information; Indigenous people have been known to answer 'yes' or agree with non-Indigenous people even if the question is not fully understood—a phenomenon known as 'gratuitous concurrence' (Liberman, 1985, p. 137). To avoid these limitations, researchers have used 'yarning', an Indigenous cultural form of storytelling and conversation, to lead unstructured data collection with Indigenous communities who speak Australian Aboriginal English (Bessarab and Ng'andu, 2010; Rodríguez Louro and Collard, 2021b).

2.2 Biases in speech recognition technologies

There is no prior work measuring performance of ASR on Australian Aboriginal English. However researchers have found that such ASR technologies often exhibit poor performance for members of other historically minoritised groups. Most work in this vein has to date been focused on African American English (AAE) speakers (Koenecke et al., 2020; Ngueajio and Washington, 2022). Martin (2022) explored how linguistic features of AAE relate to failures of ASR technologies. Tatman and Kasten (2017) found that automated captions for Black speakers were significantly less accurate than that of white speakers. Beyond AAE, Wassink et al. (2022) noted that ASR technologies exhibit poor performance for Native American, and ChicanX users in the USA. Palanica et al. (2019) found that voice assistant users with foreign accents are more likely to experience errors.

2.3 Psychological harms caused by speech recognition technologies

There is no prior work on Australian Aboriginal English speakers' experiences when using voice technologies. However prior research has found that when ASR systems make errors, some members of minoritised language communities feel that the computer doesn't like how they are speaking (Brewer et al., 2023), or that the technology isn't built for people like them (Mengesha et al., 2021). Whereas speakers of mainstream language varieties may blame the ASR system for poor performance, speakers of minoritised communities may blame themselves, and experience a drop in self esteem or an increase in self consciousness about their identity (Wenzel et al., 2023). They may feel that their language variety is being judged as illegitimate (Cunningham et al., 2024). If ASR fails to meet expectations for speakers of Australian Aboriginal English, all these impacts may be reproduced. Feelings of shame or embarrassment might also arise if speakers are self-conscious about their way of speaking being transcribed and re-contextualised into contexts where mainstream written English is the norm (Tan et al., 2021; Hutchinson et al., 2025).

3 Methods

A qualitative research approach was adopted, utilizing semi-structured interviews to gather rich, descriptive data from Aboriginal English speakers who use voice technologies. The methods were approved by the institutional review board at the authors' university. Like Kuhn et al. (2020), we also meet regularly with an Indigenous Advisory Committee for guidance. We strive to employ Indigenous Research Assistants (RAs), and to facili-

tate their engagement on the project not just as data collectors but also in actively shaping the methods and the interpretation of the data. While Author 2, the project's senior research assistant, is also a project leader, four additional RAs were also hired: Author 3, Author 4, Author 5, and Author 6. All are First Nations people of the Nyungar Nation of southwest Western Australia.

3.1 Participants

A total of 39 participants were recruited for the study. Participants were selected based on their regular use of at least one speech-enabled device or application. Requirements for eligibility included being available for in-person interviews, identifying as Aboriginal or Torres Strait Islander, and being 16 years or older. Participants were aged from 16 to 92, with a median age of 33. Over 25 First Nations were represented, with a majority from the Nyungar group, but also including Nations such as Bardi, Martu and Yamatji. Further demographic information is detailed in Appendix C.

3.2 Data Collection

Interviews were conducted in Aboriginal English, and led by the RAs. Interviews were conducted in-person at convenient locations, typically in private homes in the Perth metropolitan area. Interviews were usually conducted individually, but in some cases two participants were interviewed together, providing a relaxed environment for free-flowing discussion in keeping with cultural norms of speakers of Aboriginal English (Rodríguez Louro and Collard, 2021b; Dixon, 2017). Interviews had a median duration of 38 minutes, and were audio-recorded with participants' consent. The interview protocol (Appendix A) was based on previous studies such as Mengesha et al. (2021). Since the questions were written in formal, technical language, the RAs found that they only elicited short answers. The RAs were free to employ the protocol in ways that were appropriate for them and the participants (see pragmatic considerations in §2.1), and adapted the questions so as to elicit rich data in semi-structured interviews. Participants also provided written responses to five text prompts included on a handout (Appendix B). Each participant received $50 cash and a $100 gift card.

3.3 Data Processing and Analysis

The recorded interviews were transcribed, omitting topics of conversation that strayed from voice technologies, for the privacy and safety of participants. Then thematic analysis, as outlined by Braun and Clarke (2006), was employed to identify, and report patterns within the data, using a combination of inductive and deductive analysis. This involved a subset of the authors familiarising ourselves with the data, identifying initial themes, reviewing, defining and naming the themes, and discussing the emerging themes with the entire project team including the Indigenous RAs.

4 Findings

Our findings reveal shared patterns in the experiences of Australian Aboriginal English speakers when using ASR-powered technologies. First, technology often fails for people who speak Australian Aboriginal English. Second, speakers of Australian Aboriginal English often feel so frustrated using ASR-powered technologies that they give up on the technology. Third, speakers of Australian Aboriginal English associate ASR performance with race, reporting that the technology works better for white people and modifying their voices to sound more like non-Indigenous people.

Finding 1: Speakers of Australian Aboriginal English associate their language variety with technology failures. A recurring theme was that participants associate Aboriginal English's differences from mainstream English with poor ASR performance. Participants frequently mentioned their variety had distinct words or 'slang', which the system did not know, e.g., *'She doesn't have the vocabulary... that we have'* [F43],[1] *'we chuck in like "dardy"'*[2] [M30]. Many participants also attributed ASR failures to Aboriginal English's pronunciation differences from mainstream Australian English, e.g., *'it's probably just not catered for our accents'* [F25]. Participants often blamed ASR errors on their own language skills, such as speaking 'too fast', not speaking 'clearly', or their failure to speak mainstream English, with a few participants even questioning whether their own intelligence or level of education was the cause, e.g., *'Am I a slow learner? What's wrong with me?'* [M51], or *'maybe it's cos I didn't go to school or maybe it's because I have a low IQ'* [F25].

[1] We report direct quotes from participants using ids which combine gender and age.

[2] 'Dardy' is a Nyungar word meaning good, good-looking or cool, and is used in Nyungar Aboriginal English (Rodríguez Louro et al., 2026, p. 374).

Finding 2: Speakers of Australian Aboriginal English report frustration, and giving up. A very common theme was the frustration that participants felt when the technology failed: *'Frustration is the biggest thing there'* [M46]. Several participants reported that ASR's failures to recognise words resulted in frustration boiling over into anger, e.g., *'it's just frustrating and makes me feel a bit angry with it'* [F21]. Feelings of frustration also led many Aboriginal English speakers to give up on ASR-powered technologies, e.g., *'So I just got annoyed with it and I think I ended up giving it to someone else'* [F43], or *'I turned that part off, and now I just use it as a clock'* [F44].

Finding 3: Speakers of Australian Aboriginal English associate ASR performance with race. Another theme was that ASR-powered technology does not work well for Indigenous people in Australia. Participants frequently said that the technology worked better for white people *'because it's already catered for them and used to their language'* [F25]. Some participants thought the technology worked best for specific groups of white people, e.g., *'educated white people'* [F43] or *'posh sounding'* [F20]. Participants who did not abandon using ASR-powered technologies frequently reported changing how they speak in order to use the technology successfully. Many participants described this in terms of sounding white: *'I have to talk proper white ways... I have to change the way I speak so it understands me.'* [F45], *'So, I try to talk it white way, instead of blackfella*[3] *way, and then they'll understand.'* [M46], *'have to really talk very clear ... and try and put on my best, educated, white person voice'* [F43]. F25 described the labour of having to change how she speaks: *'it's exhausting sometimes to always keep code-switching an' acting in quotation marks like professional white or proper. Cos you just wanna like relax when you come home, I guess, and be yourself.'*

5 Discussion

This article is the first to report on the experiences of First Nations people in Australia using voice technologies. This project contributes to the growing understanding of how minoritised language communities directly connect ASR failures to a lack of inclusion in the technology of their variety (Mengesha et al., 2021), and in particular how non-white communities associate successful use of voice technologies with whiteness (Mengesha et al., 2021; Cunningham et al., 2024). Our findings also strengthen our understanding of how users code-switch to mainstream Englishes in order to use technology (Holt et al., 2024; Harrington et al., 2022). Like Wenzel et al. (2023), we found evidence that some minoritised speakers express sentiments associated with low self-esteem and blame themselves for failures in the technology.

We also found further evidence of the stressful labour that minoritised communities must perform to use voice technologies, by changing how they speak (Cunningham et al., 2024). The pressure on Indigenous people to switch to a 'white voice' when using technology echoes findings from the education and healthcare domains regarding interactions with non-Indigenous people (Harrison, 2004; Jennings et al., 2023; Fraiese, forthcoming). Because voice-powered technologies are increasingly being deployed on personal devices and in private homes, our findings reveal significant increase in the pressure to speak 'white way' into personal spaces. What is unique about our findings is that these increasing pressures are in response to a non-human agent (the technology) and, importantly, they transcend institutional contexts. While Australian Aboriginal English speakers have consistently been expected to modify their speech in institutional settings (e.g., Eades' (2006) description of a cross-examination of three Aboriginal teenage boys in a Queensland court where the defence counsel used lexical strategies to reframe what the witnesses said using mainstream English), ours is the first study to document these pressures in the home domain.

Many have called for decolonial agendas in language technology (e.g., Bird, 2020; Mohamed et al., 2020; Birhane, 2023; Schwartz, 2022). In this project, we show some ways in which this can be done in a specific context in Australia. We work closely with an advisory committee of Indigenous experts. We actively recruit Indigenous RAs and foster trust to encourage autonomy regarding how to conduct interviews in ways which make sense culturally and socially for them and their participants. We involve the Indigenous RAs in the interpretation and analysis of the data—not merely the data collection—by facilitating culturally appropriate group discussion sessions.

[3] 'Blackfella' is an Aboriginal English word meaning Aboriginal person.

6 Limitations

This research project is situated in Australia, where settler colonialism has led to the loss of many ancestral Indigenous languages, and where Indigenous people continue to be minoritised Indigenous people have frequently been viewed by research projects as subjects of research rather than partners in research, and we recognise that the legacy of unethical and settler-colonial research has contributed to distrust of research in many Indigenous communities (Smith, 2021). These layers of distrust of research may have affected who chose to take part in the interviews.

Since Aboriginal English is primarily an oral language, without standardised orthography, the project utilised the previous experiences of Author 2 and Author 1 in how to best transcribe spoken Aboriginal English in written form.

7 Ethical considerations

We recognise that university environments can often be alien or hostile for many Indigenous people in Australia. For this reason, the interviews were conducted in other places, typically private homes or public spaces. Transcripts were anonymised and screened at UNIVERSITY ANONYMIZED FOR PEER REVIEW for culturally or socially sensitive content before being shared with the non-university investigator at COMPANY ANONYMIZED FOR PEER REVIEW.

We obtained consent from all interviewees to record the interviews, and participants were also provided with a project information form. This included information about the project goals, as well as details on how Indigenous and Cultural Intellectual Property (ICIP) and Indigenous knowledge would be handled, where data will be kept, how participants can obtain a copy of their data, and how privacy will be preserved. It also provided contact details for one of the project leads.

References

Alëna Aksënova, Daan van Esch, James Flynn, and Pavel Golik. 2021. How might we create better benchmarks for speech recognition? In *Proceedings of the 1st Workshop on Benchmarking: Past, Present and Future*, pages 22–34, Online. Association for Computational Linguistics.

Dawn Bessarab and Bridget Ng'andu. 2010. Yarning about yarning as a legitimate method in Indigenous research. *International Journal of Critical Indigenous Studies*, 3:37–50.

Steven Bird. 2020. Decolonising speech and language technology. In *28th International Conference on Computational Linguistics, COLING 2020*, pages 3504–3519. Association for Computational Linguistics (ACL).

Abeba Birhane. 2023. Algorithmic colonization of Africa. *Imagining AI: How the World Sees Intelligent Machines*, page 247.

Claire Bowern. 2023. *The Oxford guide to Australian languages*. Oxford University Press.

Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101.

Robin N Brewer, Christina Harrington, and Courtney Heldreth. 2023. Envisioning equitable speech technologies for Black older adults. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 379–388.

Jay Cunningham, Su Lin Blodgett, Michael Madaio, Hal Daumé III, Christina Harrington, and Hanna Wallach. 2024. Understanding the impacts of language technologies' performance disparities on African American Language speakers. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 12826–12833.

Sally Jane Dixon. 2017. *Alyawarr children's variable present temporal reference expression in two, closely-related languages of Central Australia*. Ph.D. thesis, The Australian National University.

Diana Eades. 2006. Lexical struggle in court: Aboriginal Australians versus the state. *Journal of Sociolinguistics*, 10(2):153–180.

Diana Eades. 2012a. Communication with Aboriginal speakers of English in the legal process. *Australian Journal of Linguistics*, 32(4):473–489.

Diana Eades. 2012b. The social consequences of language ideologies in courtroom cross-examination. *Language in society*, 41(4):471–497.

Diana Eades. 2013. *Aboriginal ways of using English*. Aboriginal Studies Press.

Lucía Fraiese. forthcoming. Sociolinguistic identities by First Nations women in an Indigenous boarding school. *Journal of Linguistic Anthropology*.

Jean Harkins. 1994. *Bridging two worlds: Aboriginal English and crosscultural understanding*. University of Queensland Press.

Christina N Harrington, Radhika Garg, Amanda Woodward, and Dimitri Williams. 2022. "It's kind of like code-switching": Black older adults' experiences with a voice assistant for health information seeking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–15.

Neil Harrison. 2004. Self-recognition and well-being: Speaking Aboriginal English in healthy classrooms. *TheAustralian Journal of Indigenous Education*, 33:7–13.

Faye Holt, William Held, and Diyi Yang. 2024. Perceptions of language technology failures from South Asian English speakers. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 4067–4081.

Ben Hutchinson, Celeste Rodríguez Louro, Glenys Collard, and Ned Cooper. 2025. Designing speech technologies for Australian Aboriginal English: Opportunities, risks and participation. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, pages 108–124.

Tahir Javed, Sumanth Doddapaneni, Abhigyan Raman, Kaushal Santosh Bhogale, Gowtham Ramesh, Anoop Kunchukuttan, Pratyush Kumar, and Mitesh M Khapra. 2022. Towards building ASR systems for the next billion users. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 10813–10821.

Warren Jennings, Sonya Egert, Celestine Fisher, Sonia Renouf, Vivian Bryce, Sean Grugan, William Wang, and Deborah Askew. 2023. Better cardiac care—the patient experience—a qualitative study. *International Journal for Equity in Health*, 22(1):122.

Allison Koenecke, 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):7684–7689.

Roland Kuhn, Fineen Davis, Alain Désilets, Eric Joanis, Anna Kazantseva, Rebecca Knowles, Patrick Littell, Delaney Lothian, Aidan Pine, Caroline Running Wolf, et al. 2020. The Indigenous Languages Technology project at NRC Canada: An empowerment-oriented approach to developing language software. In *Proceedings of the 28th international conference on computational linguistics*, pages 5866–5878.

Kenneth Liberman. 1985. *Understanding interaction in Central Australia*. Routledge.

Ian G Malcolm. 2000. Aboriginal English: From contact variety to social dialect. *Processes of language contact: Studies from Australia and the South Pacific*, pages 123–144.

Joshua L Martin. 2022. *Automatic Speech Recognition Systems, Spoken Corpora, and African American Language: An Examination of Linguistic Bias and Morphosyntactic Features*. University of Florida.

Zion Mengesha, Courtney Heldreth, Michal Lahav, Juliana Sublewski, and Elyse Tuennerman. 2021. "I don't think these devices are very culturally sensitive."—Impact of automated speech recognition errors on African Americans. *Frontiers in Artificial Intelligence*, 4:725911.

Shakir Mohamed, Marie-Therese Png, and William Isaac. 2020. Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4):659–684.

Mikel K Ngueajio and Gloria Washington. 2022. Hey ASR system! Why aren't you more inclusive? Automatic speech recognition systems' bias and proposed bias mitigation techniques. A literature review. In *International conference on human-computer interaction*, pages 421–440. Springer.

Adam Palanica, Anirudh Thommandram, Andrew Lee, Michael Li, and Yan Fossat. 2019. Do you understand the words that are comin outta my mouth? voice assistant comprehension of medication names. *NPJ digital medicine*, 2(1):55.

Celeste Rodríguez Louro and Glenys Collard. 2024. The Yarning Corpus: Aboriginal English in Southwest Western Australia. *Australian Journal of Linguistics*, pages 1–17.

Celeste Rodríguez Louro and Glenys Collard. 2021a. Australian Aboriginal English: Linguistic and sociolinguistic perspectives. *Language and Linguistics Compass*, 15(5):e12415.

Celeste Rodríguez Louro and Glenys Collard. 2021b. Working together: Sociolinguistic research in urban Aboriginal Australia. *Journal of Sociolinguistics*, 25:785–807.

Celeste Rodríguez Louro and Glenys Collard. 2024. Hearing the voices: Embracing diversity in the study of language in society. *Voice and Speech Review*, pages 1–8.

Celeste Rodríguez Louro, Glenys Collard, and Madeleine Clews. 2026. Australian Aboriginal English. In Raymond Hickey, editor, *The New Cambridge History of the English Language. Volume VI: English in Africa, Asia, Australasia and the Pacific*, pages 361–384. Cambridge University Press.

Lane Schwartz. 2022. *Primum Non Nocere*: Before working with Indigenous data, the ACL must confront ongoing colonialism. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, volume 2.

Linda Tuhiwai Smith. 2021. *Decolonizing methodologies: Research and indigenous peoples*. Bloomsbury Publishing.

Kim Hua Tan, Michelle Elaine anak William Jospa, Nur-Ehsan Mohd-Said, and Mohd Mahzan Awang. 2021. Speak like a native English speaker or be judged: A scoping review. *International journal of environmental research and public health*, 18(23):12754.

Rachael Tatman and Conner Kasten. 2017. Effects of talker dialect, gender & race on accuracy of Bing speech and YouTube automatic captions. In *Interspeech*, pages 934–938.

Gari Tudor-Smith, Paul Williams, and Felicity Meakins. 2024. *Bina: First Nations languages, old and new*. La Trobe University Press.

Alicia Beckford Wassink, Cady Gansen, and Isabel Bartholomew. 2022. Uneven success: automatic speech recognition and ethnicity-related dialects. *Speech Communication*, 140:50–70.

Kimi Wenzel, Nitya Devireddy, Cam Davison, and Geoff Kaufman. 2023. Can voice assistants be microaggressors? Cross-race psychological responses to failures of automatic speech recognition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

A Interview protocol

Each of the Indigenous Research Assistants was provided with a data collection protocol and a smartphone for recording the interviews. The data collection protocol outlined the purposes of the study, the interviewee selection criteria, details on what to collect and what not to collect, payment instructions, recording instructions, and the following questions for the semi-structured interviews.

1. Where do you live?
2. Are you working now? If so, where? / What work did you use to do?
3. Are you studying? If so, what course of studies? / Did you attend Uni?
4. What is the language you speak best?
5. Do you speak other languages or varieties?
6. Do you speak a language other than Aboriginal English or 'home talk'? Do you speak it a lot?
7. How often do you use your voice to get things done on your devices? For example, how often do you use your voice to ask your phone to look something up for you? How do you use your voice with the device? Any specific examples?
8. Do you remember a time when you spoke to your phone and it didn't understand you? Can you tell me / us about it?
9. When using voice technology, how often do errors occur?
10. You mentioned that errors sometimes occur when you use voice technology. In a sentence or two, what are the main reasons you believe these failures or errors occur when using voice technology?
11. What are the main reasons you believe these errors occur when using voice technology?
12. Do you agree with the following statement? "I am the first one to buy new technology. It is important for me to explore new technologies".

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

13. Do you agree with the following statement? "I tend to buy or try out new technology much sooner than others, usually trusting my own intuitions."

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

14. Do you agree with the following statement? "I buy new technologies that have practical benefits for me; I read reviews and talk to people before purchasing."

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

15. Do you agree with the following statement? "I buy technology once most people are using it and it has been around for a while."

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

16. Do you agree with the following statement? "I am generally not interested in buying or trying out new technology."

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

17. What works well for you when it comes to voice technology (speaking to, dictating a message, or commanding your device to do something)?

18. What doesn't work well for you when it comes to voice technology (speaking to, dictating a message, or commanding your device to do something)?

19. Have you ever modified the way you talk to get different results when using voice technology?

20. In a few full sentences, tell us about when and why you modify the way you talk to get different results when using voice technology.

21. Which emotions, if any, do you feel when modifying the way you talk to get different results when using voice technology?

22. Please indicate the extent to which you agree or disagree with the following statement: "I need to modify the way I talk to get different results when using voice technology because the technology *works better for people who are not like me".

23. Who does the technology work better for?

24. Do you agree or disagree with the following statement? "I need to modify the way I talk to get different results when using voice technology because the technology doesn't understand people who come from my racial group".
    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

25. Please indicate the extent to which you agree or disagree with the following statement: "I need to modify the way I talk to get different results when using voice technology because the technology doesn't understand me due to my accent."

26. Imagine that you could speak directly to a company that designs voice technology. If you could wave a magic wand, what would you ask this company to change to improve the voice technology experience for yourself?

27. What would you ask this company to change to improve your voice technology experience, and why? Why do you think that this needs to be changed?

28. Rate your level of agreement with the following statement: "When my smart assistant speaks back to me, I want my smart assistant to sound like me."

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

29. Rate your level of agreement with the following statement: "When my smart assistant speaks back to me, it should mirror the way I speak."

    [Strongly Agree | Agree | Neutral/Not Sure | Disagree | Strongly Disagree]

B Text prompts

Text prompts for written responses
When using voice technologies, errors occur when I...
Because of these voice technology errors, I do...
These voice technology errors make me feel...
To make things better, technologists should...
I find voice technologies useful for...

Table 1: Text prompts for written responses on voice technology errors.

C Demographics

Age	Count
16–24	10
25–34	10
35–44	6
45–54	6
55–64	4
65+	2
Not provided	1

Table 2: Age of participants.

Gender	Count
Female	23
Male	15
Non-binary	0
Not provided	1

Table 3: Gender of participants.

First Nations general affiliation	Count	Notes
Aboriginal	3	Generic label.
Badimaya	2	Northern Western Australia. Mt Magnet and surrounds.
Bardi	1	Northern Western Australia. Broome and parts of the Dampier Peninsula.
Goomala	1	Northern Western Australia. Pilbara.
Jigalong	1	Northern Western Australia. Pilbara.
Karrajarri	1	Northern Western Australia. Northern Pilbara.
Larrakia	1	Northern Territory. Darwin.
Martu Wanka	1	Northern Western Australia. Sandy Desert.
Naaguja	2	Midwest Western Australia. Geraldton and surrounds.
Nyungar	24	Collective bloc from southwest Western Australia, from Geraldton to Esperance.
Wadjarri	4	Northern Western Australia. Murchison region.
Wakaya	1	Northern Territory.
Martu	1	Northern Western Australia.
Yamatji	4	Northern Western Australia.

Table 4: First Nations affiliations of participants. Each interviewee might identify with multiple First Nations, and might identify with specific groups within each Nation. In Australia, 'Aboriginal' is a general term for First Nations people, and some identify with this general term. Other spellings of 'Nyungar' include 'Noongar' and 'Nyoongar, depending on individual community members' preferences; we prefer the spelling 'Nyungar', as established in Author 2's longstanding research on Australian Aboriginal English.

Nyungar group	Count	Notes
Balladong/ Baladong/ Balardong	4	Southwest Western Australia. Northam, York, Beverley, Goomalling, Cadoux, Koorda, Wyalkatchem, Goomalling and Cunderdin. The towns of Kellerberrin and Merredin are further east, and to the south-east are the towns Bruce Rock, Narembeen, Hyden, Kondinin, Kulin, Dumbleyung, Kukerin, Lake Grace and Varley. Author 2 is from this group.
Bibbulman	1	Southwest Western Australia. Denmark, Nannup, and Manjimup.
Gnaala Karla	2	Southwest Western Australia. The Gnaala Karla Booja region refers to the Nyungar language or dialectical groups of the Binjareb/Pinjarup, Wilman, Ganeang, and Wardandi. Towards the east of the Gnaala Karla Booja region is the Balladong region and down south is the Wagyl Kaip region.
Goreng	5	Southern Western Australia. Gnowangerup and Bremer Bay area.
Mineng	2	Southern Western Australia. Albany.
Wadjuk	2	Southern Western Australia. Perth.
Wilormin/ Wirlomin	1	Southwest Western Australia.
Yuat/Yued	2	Southwest Western Australia. The Yued region includes the towns of Leeman, Jurien Bay, Cervantes, Two Rocks, Toodyay, Gingin, Calingiri, Dalwallinu, Coorow and Moora.

Table 5: Specific Nyungar affiliations of participants. Each interviewee might have multiple affiliations. In some cases, more than one spelling for the name of the group is in widespread use, depending on preferences of individual community members.

D Additional findings from interviews

Australian Aboriginal English speakers encounter ASR technologies in a wide variety of contexts. Australian Aboriginal English speakers reported using ASR technologies in a variety of locations. The most frequently mentioned explicitly was in the car, although often home—the location of the interviews—was the implied context. Other reported locations where ASR was used included at work, and on a golf course. Speakers reported using ASR technologies for a variety of use cases, the most frequently mentioned being playing music, maps/navigation, and sending text messages. Other reported use cases included calling people, issuing commands to a virtual assistant, and searching for information. Many speakers also reported encountering speaker identification technologies on the phone line for the Australian Federal Government social security agency known as Centrelink.

Australian Aboriginal English speakers modify their voices because the technology works better for people who are not like them. Some interviewers directly asked participants 'Please indicate the extent to which you agree or disagree with the following statement: "I need to modify the way I talk to get different results when using voice technology because the technology works better for people who are not like me".' 17 respondents gave positive responses ('yeah', 'yes', 'agree', or 'strongly disagree'), while 3 gave explicitly negative responses ('no' or 'nah').

Australian Aboriginal English speakers modify their voices because the technology doesn't understand people who come from their racial group. Some interviewers directly asked participants 'Do you agree or disagree with the following statement? "I need to modify the way I talk to get different results when using voice technology because the technology doesn't understand people who come from my racial group".' 13 respondents gave positive responses ('yeah', 'mm-hmm', 'agree', or 'strongly disagree'). 2 gave explicitly negative responses ('disagree' or 'nah'), while 2 more suggested alternative theories that the technology either fails to understand everyone, or that it gets confused by *'what you're trying to ask it'* (such as asking for Bindi Bindi, a song in Nyungar language).

Australian Aboriginal English speakers modify their voices because the technology doesn't understand their accent. Some interviewers directly asked participants 'Do you agree or disagree with the following statement? "I need to modify the way I talk to get different results when using voice technology because the technology doesn't understand me due to my accent.".' 13 respondents gave positive responses ('yeah', 'yes', 'agree', or 'strongly disagree'). 2 gave explicitly negative responses ('disagree or 'nah'), while 3 said they don't think of themselves as having accents.

Australian Aboriginal English speakers want more inclusive voice technologies. Some interviewers directly asked participants what they would ask for if they could speak to a technology company. Responses included *'include dialect from multiple backgrounds'*, *'a better understanding of our accent'*, *'more research and more data collection'*, *'take in consideration and maybe get like people of different backgrounds. Like Aboriginal but also people from other backgrounds'*, *'they need to actually have, you know, various Aboriginal speakers and what not using different words and getting it in their system'*, and *'be more compassionate with the words that are day to day with Aboriginal people'*. Some suggested inclusive processes, e.g., *'actually communicate to someone in them [First Nations] countries to translate it properly on their technology'*, and *'Go and visit all the language groups, and actually have people, have* [4] *actually doing it'*.

[4] 'Mob' is a word of Aboriginal English which refers to Indigenous people in Australia.

E Selected data from interviews

Finding 1: Speakers of Australian Aboriginal English associate their language variety with technology failures

'I think a lot of Aboriginal people speak fast sometimes when it's like slang or Aboriginal English' [F25].
F23 mentioned experiencing more errors when using *'my Aboriginal English where it's little bit slang'*.
'we chuck in like 'dardy" [M30].
F23 spoke of using words that are *'not a typical English word, or name'*. *'our own unique blend'* [F33].
'Aboriginal people have unique accents' [F21].
'I feel proud because I feel well at least I can be black and I can be white.' [M46].
'if I'm speaking and I'm giving a presentation to mob or what not, well I need to ... explain it the way that they understand it.' [M27].
'you know them mob talk broken English. <skip> Remote community mob. But I mean, we can talk broken English if we have to, you know.' [M46].
F43 mentioned that the technology *'doesn't know the words'* [F43].
'Siri or Google find it hard to decipher on the contact list that I may have who that might be' [F23].
F21 said of Aboriginal people: *'their voices aren't represented'* in the technology.
'not all Aboriginal peoples speak or sound the same' [F25].
'it's not just one accent, there's like… hundreds of different Aboriginal groups across Australia' [F21].
'Nyungars have like a certain way of speaking and then when you go up North they have a different accent up there as well' [F25].
'[Written prompt: When using voice technologies, errors occur when I...] speak with my natural accent' [LEH95]
'you feel like you didn't get no, um, education' [M51].
'maybe it's cos I didn't go to school' [F41].
'the way that we speak is not the same as a typical, um, person who has… been educated with the full extent of the English language.' [F23].
'I have got broken English, my English not proper and um I don't pronounce my Hs properly' [F54].
'I don't speak proper English with my h's' [F54].
'I felt like I was dumb in a way' [F21].
'I feel like I'm behind' [M51].
'maybe I'm not smart like everyone else' [F41].
'maybe it's because I have a low IQ' [F41].
'I didn't know how to present myself in a different way, I don't, like to, for somebody else to understand me, for me to get what I need' [F32].

Finding 2: Speakers of Australian Aboriginal English report frustration, and giving up

'Frustration is the biggest thing there' [M46].
'at first I get frustrated, I'm like "what the fuck?"' [F41].
'Oh, well it frustrates me' [M27].
F23 explained that when she has to repeat herself *'it just becomes really frustrating. I get a little bit, like, angry and then I'm like "oh my God. How annoying."'*.
'it should reflect your voice. And when it doesn't it's just frustrating and makes me feel a bit angry with it.' [F21].
'So I just got annoyed with it and I think I ended up giving it to someone else. And never bought another one.' [F43].
'when I got frustrated with it, I hang up. And that's all you can do.' [F44].
'Well I get angry and hang up, unna?' [F54].
'Frustrated. Hey, sometimes I'll drop the phone, smash it up, but lucky it's got a cover on.' [M46].
'She get frustrated too, yellin at that thing. And she's like "rreggedd!" And I'm like, "oh, God, she's gonna smash that iPad. I swear to God, she gonna… "' [F41].

[Written prompt: These voice technology errors make me feel...] *'Frustrated. Dumb.'* [KKS83]
[Written prompt: These voice technology errors make me feel...] *'A bit frustrated & upset, also excluded.'* [KCO04]
[Written prompt: These voice technology errors make me feel...] *'Frustrated'* [HMN70]
[Written prompt: These voice technology errors make me feel...] *'Angry'* [KEG04]
[Written prompt: These voice technology errors make me feel...] *'Annoyed'* [KTO01]
[Written prompt: These voice technology errors make me feel...] *'Angry'* [KCC91]
[Written prompt: These voice technology errors make me feel...] *'Angry'* [KBG97]
[Written prompt: These voice technology errors make me feel...] *'Frustrated at times'* [KTR73]
[Written prompt: These voice technology errors make me feel...] *'Like smashing my phone'* [KBG72]
[Written prompt: These voice technology errors make me feel...] *'Annoyed + angry + frustrated'* [LHU01]
[Written prompt: These voice technology errors make me feel...] *'Sad. Nyam. '* [LEH95]
[Written prompt: These voice technology errors make me feel...] *'Frustrated & upset because the technology doesn't understand me. '* [LAU00]
[Written prompt: These voice technology errors make me feel...] *'Frustrated'* [KPC81]
[Written prompt: These voice technology errors make me feel...] *'Frustrated and wonder why I bother trying to use them'* [KMS06]
[Written prompt: Because of these voice technology errors, I do...] *'Don't bother to use it much'* [HMN70]
[Written prompt: Because of these voice technology errors, I do...] *'Yell at her'* [KEG04]
[Written prompt: Because of these voice technology errors, I do...] *'Type it in'* [KTO01]
[Written prompt: Because of these voice technology errors, I do...] *'Go back and edit'* [KCO04]
[Written prompt: Because of these voice technology errors, I do...] *'Yell at her'* [KBG97]
[Written prompt: Because of these voice technology errors, I do...] *'Be patient then word each each syllable'* [KTR73]
[Written prompt: Because of these voice technology errors, I do...] *'Slow down or move to another room so no background noise'* [KBG72]
[Written prompt: Because of these voice technology errors, I do...] *'Repeat words'* [LHU01]
[Written prompt: Because of these voice technology errors, I do...] *'Tend not to use voice technology or I stop using it whilst I'm speaking.'* [LAU00]
[Written prompt: Because of these voice technology errors, I do...] *'Go back to typing'* [KMS06]

Finding 3: Speakers of Australian Aboriginal English associate ASR performance with race

[When asked who the technology works better for] The whiter community out there, them other ones [F41]]
[When asked who the technology works better for] Well I suppose, white people [M27]
[When asked who the technology works better for] Caucasian people [F33]
[When asked who the technology works better for] White people, especially posh sounding [F20]
[When asked who the technology works better for] I'm guessing the Americans [F56]
[When asked who the technology works better for] someone who… is probably… white Australian. Or… someone who's very educated with English [F23]
[When asked who the technology works better for] White people because it's already catered for them and used to their language [F25]
'I have to talk proper white ways... I have to change the way I speak so it understands me.' [F45].
'So, I try to talk it white way, instead of blackfella way, and then they'll understand.' [M46].
'have to really talk very clear ... and try and put on my best, educated, white person voice' [F43].
'Only when I like talk about it quotation marks like proper English will it understand straight away.' [F25].
'I had to change my slang. I had to talk, like, proper English' [F33].
'I had to… put on a voice to actually be heard' [F21].
'it's exhausting sometimes to always keep code-switching an' acting in quotation marks like professional white or proper. Cos you just wanna like relax when you come home, I guess, and be yourself.' [F25].
[When asked who the technology works better for] Educated white people [F43]

[When asked who the technology works better for] Well, obviously, oh I don't wanna sound racist, but whitefellas [F20]
[Written prompt: Because of these voice technology errors, I do...] *'Change the way I speak'* [KKS83]]
[Written prompt: Because of these voice technology errors, I do...] *'I have to adjust my speech'* [KCC91]
[Written prompt: Because of these voice technology errors, I do...] *'I can voice my accent. Speak slower.'* [LEH95]
[Written prompt: Because of these voice technology errors, I do...] *'Change the way I speak'* [KPC81]