

Towards Culturally Grounded Emotional AI via a Community-Centred Approach

Himashi Rathnayake

University of Auckland, New Zealand
hrat060@aucklanduni.ac.nz

Background and Motivation

Today, artificial intelligence (AI) systems are increasingly embedded in everyday life, shaping communication, knowledge production, and decision-making processes. This raises questions about who is represented and whose knowledge is included in AI systems. Yet current technologies support only a small subset of the world’s languages, approximately 100 out of over 7,000 (Pratap et al. 2024), leaving many Indigenous languages excluded from modern AI systems. This imbalance raises concerns about bias and unequal performance, as existing AI systems may fail to represent speakers from diverse linguistic and cultural backgrounds accurately or fairly. In speech-based systems, for example, users are often required to adapt their speech in order to be understood. This raises a critical question: do such systems also impose constraints on how emotions can be expressed?

To address this question, it is necessary to investigate how emotions are represented in emotional AI systems, defined as systems designed to recognise or respond to human emotions (Ho 2024). Our literature review of emotional AI systems identified several limitations. Firstly, most emotional AI systems rely on a limited set of emotions derived from Western psychological frameworks, typically represented using English labels (e.g., happiness, sadness, anger, fear, disgust, surprise). These emotions are assumed to be universal (Ekman 1992) and directly translated across languages without sufficient consideration of cultural variation in emotional expression. This overlooks ongoing debates about the cross-cultural variability of emotional concepts and expression (Russell 1993; Crawford 2021), and risks misrepresenting Indigenous perspectives and knowledge. Secondly, there is a limited understanding of how emotions are conceptualised and expressed across different linguistic and cultural contexts, particularly within Indigenous communities, and a lack of resources to support the development of this knowledge. Thirdly, existing emotional AI systems lack meaningful community involvement in their design and established methodologies that begin with community perspectives. Hence, these systems may not align with cultural values, Indigenous worldviews, or emotional frameworks.

This research addresses these challenges in the context of

te reo Māori, the only Indigenous language of New Zealand. In this paper, “Māori” refers to the Indigenous people, while “te reo Māori” refers specifically to the language. Te reo Māori remains under-represented in AI technologies, and there are very limited studies on Māori emotions. Prior research has shown that Māori and New Zealand European may exhibit different emotional patterns across social situations (Banks 1996). Moreover, nuances of Māori emotions can be lost when translated to English (Elder 2020). Thus, directly applying English-defined emotion categories, or their closest translations, may be insufficient for representing emotional expressions in this language. These limitations highlight the need for culturally grounded approaches to emotional AI that respect Indigenous knowledge systems.

Accordingly, this research addresses the following question: *How can emotional AI systems be co-developed with Indigenous communities in ways that respect cultural knowledge, support data sovereignty, and avoid imposing external frameworks?* Specifically, this work investigates this question in the context of Māori speech emotion recognition. We develop a community-centred methodology for identifying culturally relevant emotions, building an emotional speech corpus, and developing a culturally grounded emotion recognition system. This approach contributes to more culturally inclusive and ethically grounded AI, advancing discussions on fairness for Indigenous languages.

This research forms the basis of the author’s doctoral thesis, building on an initial idea conceptualised with Māori technology developers at Te Hiku Media (2026a). It is a collaborative effort involving a transdisciplinary research team, including language experts, half of whom identify as Māori. All phases of this study are discussed within the team, and feedback is continuously incorporated to ensure that culturally appropriate practices are upheld at every stage.

Methodology

Our community-centred methodology is outlined below.

1. Community media collection: There was no annotated te reo Māori emotional speech database available to start exploring emotions, as is the case for many Indigenous languages. To address this gap, we collected emotional speech data from community media sources, including Te Hiku Media radio recordings and Māori television programmes, with appropriate permissions.

2. Online questionnaire: We designed a questionnaire to obtain community feedback on emotions conveyed in the community media we collected. Participants were asked to listen to selected audio clips and provide free-form descriptions of the emotions they perceived. We designed the questionnaire to be bilingual, as all Māori speakers are bilingual in English and te reo Māori. This reflects the impact of te reo Māori as a revitalised language, in which speakers may use English to express emotions while still conveying a culturally Māori emotional framework. From these responses, we extracted over 200 emotion terms.

3. Focus group: The terms were further categorised through focus group discussions with community members without losing cultural nuances. This process yielded 16 emotion categories, addressing a key knowledge gap in Māori emotion categorisation identified at the beginning of this research. Some of these emotions overlap with the commonly used presumptive universal emotions (e.g., harikoa - happy and pōuri - sad). However, we also found emotions that reflect unique cultural concepts, such as pai (good, well), which lacks a direct English equivalent, and aroha (love), a culturally rich concept often absent from mainstream emotion research, highlighting important cultural distinctions. This community-derived emotion set provides a culturally grounded foundation for future emotional AI, while also contributing to broader cultural understanding.

4. Emotional speech recording: Then we developed Pā-Kakare, the first Māori emotional speech corpus, which was needed for further understanding how Māori emotions are expressed. Unlike existing emotional speech resources that rely on presumed universal emotions, our corpus is grounded in the community-derived 16 emotion categories, ensuring cultural relevance. The corpus consists of 3,840 high-quality recordings from four professional Māori actors. To uphold principles of data sovereignty and kaitiakitanga (guardianship), the dataset is treated as a community resource rather than as researcher-owned data. Its use is governed by a Kaitiakitanga license (Te Hiku Media 2026b), ensuring culturally appropriate access, use, and protection.

5. Acoustic feature analysis: As the next step, we investigated whether emotions expressed in Māori speech exhibit patterns similar to commonly assumed universal emotion patterns. We analysed a range of acoustic features, including pitch, loudness, speech rate, vocal tract characteristics, and glottal activity. The results indicate that while some patterns of emotional expression are consistent across cultures, others differ, highlighting both cross-cultural similarities and culturally specific modes of emotional expression.

6. Technology development: Our final objective is to examine which approaches are most effective for developing an emotion recognition system tailored to Māori. Using the Pā-Kakare corpus and the extracted acoustic features, we will evaluate a range of existing approaches. Particularly, we will evaluate whether recent emotion representation learning models, such as emotion2vec (Ma et al. 2024), which are trained on large-scale datasets based on commonly assumed emotions and are often presented as universal representations, can effectively differentiate Māori emotions.

Contributions and Significance

This work contributes: (1) a community-derived set of Māori emotion categories for meaningful and culturally grounded emotion analysis, challenging the direct transfer of Western emotion frameworks; (2) a novel, replicable, community-centred methodology for emotional AI development that prioritises cultural values, data sovereignty, and ethical co-design, centring community knowledge rather than imposed frameworks; (3) empirical findings showing that while some patterns of emotional expression are shared across cultures, others are culturally specific, highlighting the limitations of assuming universal emotion representations; and (4) an evaluation of whether existing emotion recognition approaches can generalise to Indigenous language contexts such as Māori. More broadly, this research highlights the need to critically examine whether existing methodologies adequately represent diverse communities, and emphasises the importance of grounding AI development in Indigenous worldviews. It further provides a practical example of how community perspectives can be incorporated into real-world AI system development, contributing to the development of more ethical, inclusive, and socially responsible AI.

Acknowledgments

I sincerely thank my supervisors, Dr Jesin James, Dr Ake Nicholas, Dr Gianna Leoni, Professor Catherine Watson, and Associate Professor Peter Keegan, for their guidance and support. I also wish to acknowledge the financial support of New Zealand's Science for Technological Innovation National Science Challenge, Te Hiku Media and the University of Auckland.

References

- Banks, C. 1996. *A comparison study of Maori and Pakeha emotional reactions to social situations that involve whakamaa*. Ph.D. thesis, Massey University.
- Crawford, K. 2021. Artificial intelligence is misreading human emotion. *The Atlantic*, 27.
- Ekman, P. 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4): 169–200.
- Elder, H. 2020. *Aroha: Maori wisdom for a contented life lived in harmony with our planet*. Random House.
- Ho, M.-T. 2024. What is a Turing test for emotional AI? *AI & SOCIETY*, 39(3): 1445–1446.
- Ma, Z.; Zheng, Z.; Ye, J.; Li, J.; Gao, Z.; Zhang, S.; and Chen, X. 2024. emotion2vec: Self-supervised pre-training for speech emotion representation. In *Findings of the Association for Computational Linguistics*, 15747–15760.
- Pratap, V.; Tjandra, A.; Shi, B.; Tomasello, P.; Babu, A.; Kundu, S.; Elkahky, A.; Ni, Z.; Vyas, A.; Fazel-Zarandi, M.; et al. 2024. Scaling speech technology to 1,000+ languages. *Journal of Machine Learning Research*, 25(97): 1–52.
- Russell, J. A. 1993. Forced-choice response format in the study of facial expression. *Motivation and Emotion*, 17.
- Te Hiku Media. 2026a. Te Hiku Media. <https://tehiku.nz/>.
- Te Hiku Media. 2026b. Whare Kōrero Kaitiakitanga License. <https://xn--wharekrero-v3b.nz/kaitiakitanga>.