Pak. j. life soc. Sci. (2025), 23(2): 967-984 E-ISSN: 2221-7630;P-ISSN: 1727-4915

Pakistan Journal of Life and Social Sciences

www.pjlss.edu.pk

https://doi.org/10.57239/PJLSS-2025-23.2.00226

RESEARCH ARTICLE

Artificial Intelligence and Punjabi Language Revitalization: Opportunities, Challenges, and Policy Interventions

Prof. Dr. Mujahida Butt[1], Dr. Abdul Rouf[2], Dr. Muhammad Irfan ul Haq[3], Dr. Kalyan Singh Kalyan[4], Dr. Asma Ghulam Rasool[5], Muhammad Asim Khan[6]

[1]Chairperson Department of Punjabi, Lahore College for Women University, Lahore, Pakistan
[2]Department of Pakistani Languages, National University of Modern Languages, Islamabad Pakistan 3
[3]Assistant Professor, Institute of Punjabi and Cultural Studies, University of the Punjab, Lahore Pakistan
[4]Assistant Professor, Government College University, Lahore Pakistan
[5]Assistant Professor, Department of Punjabi, Government College University, Faisalabad Pakistan
[6]Department of Applied Linguistics, Government College University Faisalabad, Pakistan

ARTICLE INFO

Received: Aug 5, 2025

Accepted: Sep 25, 2025

Keywords

Punjabi Language Revitalization
Artificial Intelligence (Ai)
Natural Language Processing (Nlp)
Speech Recognition
Machine Translation
Digital Linguistic Justice
Low-Resource Languages
Ai Bias
Linguistic Data Scarcity
Language Policy
Computational Linguistics

***Corresponding Author:**

asmahumayun63@gmail.com

ABSTRACT

This study has examined the use of Artificial Intelligence (AI) in the process of Punjabi language revival. This research was meant to ascertain how effective and limited AI driven technologies were in recovering Punjabi community with the shortfall of data, the neglect of politics, and digital exclusion. Fishman's (1991) Reversing Language Shift (RLS) model, Technological Determinism Theory (McLuhan, 1964), and Digital Linguistic Justice (Bender et al., 2021) have guided the study which brings a critical and multiple dimensional approach on the impact of AI on Punjabi. Qualitative research design has been used. Data collection method through semi structured interviews, and the analysis of data is done by using thematic analysis to find key patterns and challenges in AI based Punjabi preservation. Results of the study have shown that while AI has great prospects for the revitalization of the Punjabi language, it has not been effective due to the lack of structured linguistic data, dialectal complexities, AI bias and insufficient policy interventions. Insufficient Punjabi annotated datasets, limited adaptability of NLP and unreliable speech recognition models have placed Punjabi out of the scope of AI applications. However, Punjabi's morphological complexity, patterns of code switching and script variations (Shahmukhi and Gurmukhi) have only thrown fuel to the fire for AI training. Lack of state-backed AI research funding and policies have also been acknowledged by policymakers while the community activists have highlighted that Punjabi is still poorly funded and counted in the formal institutions. This study has increased to the fields of AI in linguistics and language policy while contributing to the field of digital language preservation.

INTRODUCTION

Languages however, have nothing to do with the mere tools of communication and they are also a carrier of cultural heritage, identity and memory (Crystal, 2000). Yet globalization, urbanization, socio-political factors have led to the decline of many indigenous languages, such as Punjabi, especially in Pakistan (Rahman, 2011). Although the most widely spoken native language in Pakistan, Punjabi is systematically marginalized in education, governance and in digital spaces (Ayres, 2009). With Urdu as the language of the nation and English as that of competition, it was to Punjabi which the erosion has occurred in public and institutional domains (Shackle, 2003). Over the past few years there has been a realization that Artificial Intelligence (AI) may be utilized for the preservation and revitalization of language. New possibilities for documenting and promoting endangered languages are technologies with the help of which AI already operates – namely Natural Language Processing (NLP), machine translation and automated speech recognition (Bird, 2020).

There is however a near chaotic lack of exploration of where and how AI can be effectively utilized to revive Punjabi.

The decline of Punjabi in Pakistan is very much enmeshed in socio-political and socio-economic structures. Moreover, unlike Sindhi that has official status and government support, Punjabi is never taught in schools and is not present on official domains (Mansoor, 2005). Therefore, younger generations tend to move towards Urdu and English which is causing its attrition. According to Fishman (1991), one of the elements trying to reverse language shift involves institutional support. While it has been thus in the case of Punjabi, the absence of government intervention has left the language increasingly vulnerable. Reverse of this trend could be possible if Punjabi was made more accessible and usable in digital spaces by AI driven Punjabi language tools.

Artificial Intelligence (AI) can become a useful tool in the language revitalization project, in the case of the Punjabi language that has previously suffered the consequences of marginalization in Pakistan. Although Punjab is predominantly a speaking Punjabi language, there has been a long history of prohibition of the use of Punjabi language in terms of schooling, administration, centralization and technology because of political, social and ideology reasons. Researchers have explained that this exclusion is backed by power arrangements in rural Punjab (Zulfiqar et al., 2024), institutionalized language ranking with an emphasis on Urdu (Hussain et al., 2024) as well as historical language policy development that has side-lined regional languages (Butt et al., 2024). There is an urgent necessity to investigate how to include AI technologies, including speech recognition, machine translation, language learning aids, etc., to help the Punjabi language revitalization. Nevertheless, equal attention should be with sociopolitical awareness and corresponding global and inclusive policy intervention in order to achieve impact.

Applications of AI for automatic speech recognition (ASR), text to speech (TTS) systems, and neural machine translation (NMT) bring success to decline languages such as Hawaiian, Maori and Welsh (Jones and Hammarström, 2021). For example, Google has substantially aided Maori language digital engagement via Google's support for Maori through AI driven language processing tools (Keegan, Mato, & Ruru, 2015). Similarly, AI can help Punjabi's revitalization by growing its presence in online platforms, social media and in automated translation systems. The most significant problem is desperate lack of data. Lack of Punjabi's digital corpus and a need for large datasets for training, AI models are yet to be sufficiently trained. (Hellwig 2018) However, AI's application in the Punjabi language revitalization remains limited due to the lack of structured and heterogeneous data sets.

At the same time, there are ethical and socio-political issues related to the role of AI in language preservation. AI can support documentation, digital engagement, but not for the total language policies, intergenerational transmission, or community based work (O'Regan, 2022). Along with this, there necessarily needs to be an endeavor to promote literacy, cultural pride and educational reforms while implementing AI for Punjabi revitalization. In addition, the usage of the Western AI technology is also problematic, as it carries risk of linguistic bias and digital colonialism, where the AI models will primarily rely on resource rich languages and neglect the others, having limited data (Bender et al., 2021). Without critical interventions, AI can do the job at failing to recognize the problems existing and reinforcing it instead.

This study is a critical evaluation of the role of AI in Punjabi language revitalization on the basis of effectiveness, limitations and socio-political implications of it. This research analyzes existing AI driven language preservation efforts and considers their viability toward Punjabi by learning whether AI can be a useful tool to reverse Punjabi's decline. The study also notes that there is a need to incorporate AI tools with the community's efforts in order to achieve sustainable language preservation.

Significance of the Study

The study holds vital importance because it conducts an analysis of Artificial Intelligence (AI) potential for revitalizing Punjabi which remains marginalized in Pakistan despite wide usage. The research examines technology tools built with Natural Language Processing (NLP) that include speech recognition and machine translation features to find technological solutions that preserve Punjabi language and enable digital involvement and transmission of knowledge to future generations. The research exposes both the lack of language data which blocks AI performance as

well as digital foreign control and social-political obstacles that obstruct language revival efforts. The research results will benefit linguists along with policymakers and educators and AI researchers as they provide a framework about using AI solutions together with community-led approaches to sustain language preservation. The study takes part in advancing research about AI technologies and diversity while showing the value of digital tools that serve the needs of minority languages.

Problem Statement

The native language of Pakistan with the largest speaker base faces official discrimination within educational institutions and public administration and digital environments thus leading to gradual institutional use reduction. Government indifference to the language and widespread preference for Urdu and English has yielded minimal resources for documenting and studying Punjabi as well as digital resources that promote its reach. Artificial Intelligence has appeared as an instrument to save endangered languages by using its capabilities in speech recognition and natural language processing (NLP) and automated translation. AI has minimal established effectiveness in Punjabi language revitalization because structured linguistic data is scarce along with the dominance of AI learning models from resource-rich languages. AI deserves thorough evaluation before Punjabi can avoid digital isolation from future technological developments because this assessment will determine its exclusion in the digital era. This research examines how artificial intelligence (AI) assists Punjabi language revitalization by studying the barriers and ethical consequences and practical opportunities regarding AI-based solutions to halt language reduction.

Research Questions

What socio-political and ethical challenges arise in using AI for Punjabi language revitalization, and how can these challenges be addressed to ensure digital linguistic justice?

To what extent can AI-driven technologies (e.g., NLP, speech recognition, machine translation) contribute to the revitalization of Punjabi, and what are their limitations in addressing language decline?

How does the lack of structured linguistic data and digital resources for Punjabi impact the effectiveness of AI-based language preservation efforts?

LITERATURE REVIEW

Research about using Artificial Intelligence (AI) for language revitalization activities has become increasingly significant throughout the last couple of years. This review evaluates AI applications as used in endangered language preservation particularly within the Punjabi language context. This review examines digital cases and technological improvements to understand both the advantages and the obstacles of using AI technologies for Punjabi revitalization efforts.

Theoretical Framework

The framework of this investigation depends on Fishman's (1991) Reversing Language Shift (RLS) Model, Technological Determinism Theory (McLuhan, 1964), and Digital Linguistic Justice (Bender et al., 2021) to analyze Artificial Intelligence (AI) contributions to Punjabi language preservation. These theories serve together to create a framework which explains how AI-driven technologies maintain linguistic survival and understand both political-social dynamics behind language marginalization and ethical challenges related to AI-based language growth.

Fishman's (1991) Reversing Language Shift (RLS) Model

The RLS Model developed by Fishman puts forth essential principles for language revitalization which combine youth-to-youth language transmission with institutional backing. Fishman (1991) explains that language shift emerges when present-day youths skip learning their native language within household and communal spaces which results in the progressive replacement of their heritage language with dominant linguistic forms. The use of Punjabi in Pakistan is limited to casual domains because official institutions choose Urdu and English over Punjabi according to Rahman (2011). RLS establishes that Punjabi revival depends on systematic inclusion throughout educational teaching content and across media outlets and digital platforms. The adoption of AI tools like NLP with ASR and NMT enables direct Punjabi access using automated text-to-speech software and virtual helpers as well as machine-produced Punjabi material (Bird, 2020). The study evaluates

both policy backing and community collaboration since AI remains insufficient for reversing language shift.

Technological Determinism Theory (McLuhan, 1964)

According to McLuhan's Technological Determinism Theory technological progression affects social organization within societies particularly concerning language utilization and communication networks. The perspective serves essential in analyzing AI language preservation because it demonstrates how new technologies strengthen linguistic dominance or give underrepresented languages increased exposure (Postman, 1992). AI has proven its capability to work with languages at scale thus presenting an opportunity to make Punjabi more accessible for digital communication spaces that mostly use Urdu and English. Technological determinism faces common criticism because it stresses technology as the primary force of change above social-economic and political determinants which impact language usage (Feenberg, 2010). The success of Punjabi revitalization through AI depends on three main factors: available linguistic data and proper government policies in addition to widespread community support. Digital linguistic colonialism becomes likely when AI tools fail to achieve inclusive design because the result benefits resourceful languages but marginalizes underrepresented languages according to Bender and colleagues (2021). This research analyzes how equally Punjabi AI solutions are being developed while also determining if they reproduce current linguistic barriers in the field.

Digital Linguistic Justice and AI Ethics

Digital Linguistic Justice as a concept focuses on both ethical matters and social politics of language technology by advocating for AI systems to advance linguistic inclusion over digital inequality conditions (Bender et al., 2021). Language marginalization of Punjabi occurs from state policies and colonial patterns together with digital world domination hierarchies (Ayres, 2009). AI applications need extensive well-documented databases yet the underdevelopment of Punjabi as a language puts it in a difficult position regarding AI adoption (Hellwig, 2018). The application of AI to language conservation experiences problems because machine learning models show biased behavior when they prefer dominant languages which have big digital databases thus excluding less spoken languages according to Hovy & Spruit (2016). This research examines how current AI systems handle Punjabi linguistic properties while also determining if these systems cause non-dominate language groups to remain excluded from the digital domain.

ALA models learn primarily from corpora that originate from Global North regions thus making queries about Artificial Intelligence-generated Punjabi text doubtful regarding its cultural accuracy (Blodgett, Barocas, Daumé, & Wallach, 2020). AI-generated content in Punjabi which adopts Western-language structures might damage the natural native grammar standards and distinct dialect rules and cultural communicational methods. This paper examines why community members who identify as linguists along with educators and speakers must actively develop AI datasets and training models because this approach enables AI to empower local languages instead of creating standardized digital content.

The chosen theoretical framework justifies its use because it enables comprehensive study of AI functions in revitalizing the Punjabi language. Through Fishman's RLS Model researchers identify that the use of AI alone does not suffice to stop language decline since it requires institutional backing and community involvement. As part of its analysis Technological Determinism examines AI's influence on language behavior patterns and explains that technologies evolve through social-political factors although they are not value-neutral. Digital Linguistic Justice brings together AI ethics along with its analysis of discrimination and online accessibility along with technological exclusion in digital spaces. Through integration with these theories researchers can perform an advanced evaluation of AI-based Punjabi revitalization effectiveness while evaluating technical elements alongside social and political and ethical factors. The research takes a unified tactic to bridge technological developments with linguistic equality through diverse authentic community-led language preservation methods which rely on computer programming.

AI in Language Revitalization: Global Perspectives

The preservation and revitalization of different indigenous languages relate to extensive use of AI technologies. Te Hiku Media in New Zealand created an automatic speech recognition model to

transcribe Te Reo Māori while reaching 92% accuracy. The development of the successful language model became possible because of extensive archival footage and community-submitted audio input during the training process which emphasizes the need for strong linguistic data paired with community engagement in AI language-based projects (Jones, 2024). Within Canada First Voices operates as an open-source web platform which lets Indigenous communities document and promote their native languages. The First Voices platform allows communities to keep their linguistic resources under their control through the process of uploading alphabets and cultural presentations and audio recordings (FirstVoices, 2023).

Indonesia's Ministry of Education, Culture, Research and Technology works together with university institutions and local residents to gather data for large language model creation. The initiative works towards safeguarding approximately 700 native languages even with the obstacles presented by insufficient documentation and the lack of agreed-upon writing conventions (Aziz, 2024). Proper success in language revitalization recovery projects requires abundant linguistic data together with community involvement and cultural conservation activities.

The Punjabi Language: Current Status and Challenges

The widespread nature of Punjabi language in Pakistan does not protect the language from severe obstacles that threaten its continued survival and rising visibility. The language faces marginalization when official institutions choose to promote Urdu and English over Punjabi in their public settings. Street areas along urban regions show reduced Punjabi usage largely due to discrimination thus creating feelings of linguistic inferiority in the speakers (Hussain et al., 2024). The University of British Columbia launched 'Punjabi in BC' as an initiative to document people responsible for promoting Punjabi usage in British Columbia. Through its collection of oral history, the project maintains great value as it depicts Punjabi language and cultural practices existing throughout diaspora communities (Abdullah, 2023). The lack of institutional backing combined with the absence of Punjabi in official sectors prevents revitalization progress in Pakistan today. The negative social perception of Punjabi speaks as a factor which intensifies the language protection problem by driving young people away from learning and using it (Hussain et al., 2024).

Potential of AI in Punjabi Language Revitalization

The utilization of AI technology for Punjabi revitalization creates both possibility and difficulty during this process. The integration of natural language processing and machine translation and speech recognition within AI technologies permits Punjabi digital resource and educational tool and accessible content development. AI-powered platforms would help generate extensive Punjabi language corpora so developers can produce learning applications with automated translation capabilities.

AI proves effective in these situations provided a number of things remain in place. Training AI models depends on substantial amounts of top-quality linguistic information. There is a major obstacle to creating efficient artificial intelligence applications for Punjabi due to the limited availability of digital versions of Punjabi texts and sound files. For a language revitalization initiative to succeed the local community needs to actively participate. For sustainable initiatives the participation of Punjabi speakers remains essential so they can both produce content with AI applications and gather data and support tool use. Government institutions together with educational establishments take central positions to maintain language survival. The process of Punjabi language reinforcement requires formal state support through educational curriculum integration and funding investments in technology development. AI applications need comprehensive knowledge about Punjabi cultural specificities and its multiple dialects because this knowledge ensures native speakers accept the application. Proper solutions need to be developed to resolve existing issues because insufficient resolution will prevent AI initiatives from achieving meaningful gains for Punjabi conservation.

The implementation of Artificial Intelligence helps create new opportunities for protecting Punjabi as well as other endangered languages. Such initiatives succeed when linguistic data is complete and when communities fully participate and receive institutional support together with technology that respects cultural principles in its development. A partnership between modern technology and local

community action aims to guarantee the digital survival of Punjabi language as an endangered language.

Different research work demonstrates how Artificial Intelligence (AI) technology enhances endangered language preservation by using Natural Language Processing (NLP) and speech recognition with machine translation systems (Jones 2024; FirstVoices 2023). AI-based language preservation projects among indigenous communities show success in their work with Te Reo Māori in New Zealand and Hawaiian in USA and various Canadian Indigenous languages (Aziz, 2024). Thousands of successful AI models exist because they depend on massive linguistic datasets supported by government policies and institutional budgets - factors that Punjabi in Pakistan significantly lacks. Extensive research has shown Punjabi faces educational and governmental along with digital space marginalization (Hussain, Khan, & Khan, 2024) yet little is known about using AI for reviving the language. Research about Punjabi language decline by Minhas and Salawu (2024) explores societal and cultural limitations of the language but avoids studying AI technology as a resource that joins traditional preservation initiatives with digital advancement. AI research in Pakistan concentrates primarily on economic medical and security domains (Abdullah, 2023) while skipping the potential of AI-driven linguistic preservation which leaves Punjabi at a disadvantage in the current AI age. Organizations face a fundamental shortage of structured linguistic data for AI training because such data serves as a core requirement for developing NLP models. Punjabi stands in stark contrast to Māori, Welsh and Icelandic because government-backed AI initiatives are absent while digital corpora remain unavailable for standardization at the national level thus making the language scarcely represented within global AI-driven language models (Jones 2024).

The processing of Punjabi phonetic and syntactic structures by AI models becomes challenging when dealing with unannotated datasets which results in poor performance for both machine translation and speech recognition systems (Aziz, 2024). AI ethics scholars highlight digital linguistic justice as essential (Bender et al., 2021) yet no study exists regarding potential ethical dangers when applying AI technology for Punjabi revitalization and the risk factors associated with AI bias and Western-created NLP tools that may distort Punjabi cultural diversity and its dialectal variants (Blodgett, Barocas, Daumé, & Wallach, 2020). The existing research gaps in this field exhibit a clear demand for more comprehensive investigations about Punjabi language revival through AI while identifying related policy restrictions and developing ethical methods to deploy AI solutions.

METHODOLOGY

The research uses qualitative methodology to thoroughly study Artificial Intelligence (AI) contributions to Punjabi language renewal efforts. Qualitative research methods provide the most appropriate approach to study the intricate situation of language decline and AI applications while investigating social and political factors affecting preservation efforts because this study focuses on exploratory research. AI-driven language revitalization receives detailed evaluation in this study by using interviews together with document analysis and thematic analysis to obtain descriptive data.

Research Design

The investigation implements multiple case studies to gain comprehensive knowledge about how AI helps language revitalization efforts between separate situations. This design enables researchers to study effective AI-based maintenance of Māori, Hawaiian and Welsh languages compared to Punjabi language preservation challenges. The study examines multiple cases to discover specific traits, barriers and successful approaches which affect Punjabi language preservation and revitalization. The research adopts an interpretive approach to understand how Punjabi-speaking communities receive and adapt to AI technologies as well as their opposition towards such technology adoption.

Sample and Participants

Researchers have chosen the participants using purposive sampling because they have expertise in AI-driven language preservation or in Punjabi language policy. A combination of linguists along with artificial intelligence researchers and policymakers and educators and community activists forms the studied group because they work in both language preservation and digital linguistics fields. A total of 25 participants will be chosen from combined academic institutions together with AI development firms and language advocacy groups and cultural organizations that specialize in

Punjabi linguistic matters. The research sample will include community members from AI programs which effectively preserved native languages so researchers can learn from their methods.

Data Collection Method

The study employs semi-structured interviews as the primary data collection method. The researcher will conduct expert interviews with professionals who specialize in AI technology together with language policy and Punjabi linguistics to obtain knowledge about the real-world potential as well as barriers to applying AI for Punjabi restoration work. The interview examination will contemplate essential matters including data attainability with technologies alongside moral issues as well as regulatory modifications. The research design includes open-ended interview questions that enable participants to share their personal experiences coupled with their views and suggestions regarding using AI for Punjabi language preservation.

Data Analysis Method

We will analyze our gathered data through thematic analysis which embraces Braun and Clarke's (2006) six-step method beginning with familiarization of data and proceeding to generating initial codes and searching for themes and reviewing themes and defining and naming themes then producing the final report. This method empowers researchers to build a stepwise system which recognizably detects the primary themes stemming from interview records and study documents. The researcher will utilize NVivo software to run thematic coding and effectively handle and sort the data for analysis. The study examines regular patterns and conflicting perspectives about AI's contribution to Punjabi language revitalization. The research will extensively investigate multiple concepts starting from technological capability, limited linguistic data availability alongside community engagement as well as digital fairness and artificial intelligence biases. The study will receive a critical evaluation based on existing research about AI and language preservation to establish new valuable insights for academic fields and policy-making frameworks.

Purposefully employing qualitative multiple case studies allows this research to deliver extensive details about using AI for Punjabi language revival. Document analysis and semi-structured interviews were utilized to study the topic from multiple dimensions and thematic analysis structured the interpretation of research findings. The study depends on qualitative methodology to understand the intricate relationship between AI technology and efforts to protect Punjabi language.

Data Analysis

Researchers from various fields such as AI science, linguistics, policy, community advocacy, and technology development provide detailed insights about the function of Artificial Intelligence (AI) in Punjabi language revitalization. Digital sustainability of Punjabi faces both support and hindrance through three main AI-based technologies known as Natural Language Processing (NLP) and speech recognition and machine translation which operate within linguistic and socio-political and technological environments. This analysis uses qualitative data thematic examination to study significant barriers such as scarce data availability and language dialect difficulty while investigating AI bias and inactive policies and insufficient public participation because these conditions affect the potential of AI-based language protection. The interpretations of this research draw on Fishman's (1991) Reversing Language Shift (RLS) Model and the Technological Determinism Theory (McLuhan, 1964) with Digital Linguistic Justice (Bender et al., 2021) to analyze the technology-policy-community interactions. By studying stakeholder groups independently researchers produce a comprehensive view of AI's effects on Punjabi language maintenance which enables both the discovery of obstacles together with evidence-led methods to introduce AI to language revitalization initiatives.

Table 1. AI Researchers Data analysis

Participant Number	Participant Role	Key Theme	Key Insights	Recommendations
1	AI Researcher	Data Availability	Limited Punjabi language corpora hinder NLP model training. Lack of structured, annotated data makes AI-based linguistic models ineffective.	Develop collaborative, community-driven datasets and integrate speech-to-text technology.

2	AI Researcher	Machine Learning Models	Absence of pretrained AI models for Punjabi reduces efficiency in NLP tasks. Most AI models are optimized for high-resource languages.	Invest in creating domain-specific Punjabi NLP models for better AI-based language processing.
3	AI Researcher	AI Bias	AI models favor dominant languages like English and Mandarin, excluding Punjabi in global NLP research. Leads to weaker AI performance.	Ensure inclusive AI development by designing ethical AI models that incorporate Punjabi representation.
4	AI Researcher	Speech Recognition	Punjabi speech recognition lacks accuracy due to dialectal variations and the absence of diverse training datasets.	Develop AI models that account for regional dialectal variations to improve speech recognition.
5	AI Researcher	AI Resource Allocation	Developing AI for Punjabi requires significant computational resources and funding, which are scarce. No government-backed AI initiatives exist.	Encourage AI research funding for low-resource languages and promote open-source Punjabi AI projects.

The analysis of AI researchers reveals primary challenges and methods to use Artificial Intelligence (AI) in Punjabi language revitalization which match both the theoretical foundation and survey results and investigation inquiries. The limited quantity of available Punjabi linguistic data (Participant 1) documented in the table indicates that Natural Language Processing development faces limitations which substantiates the data availability concerns expressed in the questionnaire as per Fishman's (1991) Reversing Language Shift (RLS) Model. The missing pertained Punjabi models described by Participant 2 aligns with Technological Determinism Theory (McLuhan, 1964) because they demonstrate how technology can revolutionize operations but such systems fail due to inadequate resources. This illustrates the necessity of building domain-specific Natural Language Processing solutions to enhance language processing performance.

The third participant highlighted AI bias as an imperative issue compatible with Digital Linguistic Justice (Bender et al., 2021) that demonstrates how English and Mandarin languages impact AI outputs while marginalizing Punjabi as an under resourced language which demands inclusive AI frameworks that respect fairness. Additional findings through the research indicate that speech recognition failures (Participant 4) highlight the requirement for dialect-specific AI solutions because natural variations across Punjabi-speaking regions are a focus point from linguistic questionnaire surveys regarding Punjabi morphological characteristics and code-switching patterns. The limited availability of research funds for AI appears as a major challenge according to Participant 5 because policymakers must focus on this issue after reading their questionnaire responses about government-supported AI projects. Multiple findings from this research demonstrate that Punjabi needs structured linguistic datasets together with ethical AI transformation and government support and open-source AI projects to maintain its survival in the digital AI landscape.

Table 2. Linguists Data Analysis

Participant Number	**Participant Role**	**Key Theme**	**Key Insights**	**Recommendations**
6	Linguist	Technological Constraints	AI models lack robust Punjabi syntax recognition, leading to poor NLP performance.	Enhance AI model adaptability for dialectal variations and invest in localized NLP tools.
7	Linguist	Linguistic Structure	Punjabi's rich morphology and inflectional variations make AI training complex.	Develop AI language models that capture Punjabi's grammatical complexity.

8	Linguist	Lexical Complexity	Lexical gaps in AI dictionaries prevent accurate translation between Punjabi and other languages.	Expand lexical databases for Punjabi to improve machine translation accuracy.
9	Linguist	Code-Switching	Frequent code-switching between Punjabi, Urdu, and English complicates AI model development.	Train AI models to recognize and process code-switching in multilingual contexts.
10	Linguist	Linguistic Identity	Punjabi is not formally standardized, making AI-based linguistic analysis challenging.	Encourage linguistic research to create standardized Punjabi NLP resources.

The analysis conducted by linguists reveals key language obstacles that compromise the successful application of Punjabi language revitalization through AI methods while supporting Fishman's (1991) Reversing Language Shift (RLS) Model in combination with Technological Determinism Theory (McLuhan, 1964) and Digital Linguistic Justice (Bender et al., 2021) and questionnaire findings. Existing AI models face technological limitations that hamper their ability to understand Punjabi syntax according to Participant 6 which produces unsatisfactory NLP performance (Participant 6). This confirms concerns expressed in the questionnaire about AI systems coping with Punjabi language characteristics. The linguistic structure challenges facing Punjabi language emphasize its different inflections and morphological complexity as Participant 7 demonstrates suggesting the need to develop specialized AI models that respect Punjabi linguistic features as noted in discussions on AI language representation.

The issue of lexical complexity (Participant 8) highlights translation mistakes that stem from insufficient Punjabi database content which confirms the questionnaire research on AI machine translation hurdles and shows why adding more database entries improves translation precision. Multiple languages (Punjabi, Urdu and English) used interdependently by Participant 9 create a substantial challenge for AI systems because of their difficulty in detecting multilingual language changes yet call for complex NLP training. The questionnaire stresses linguistic research importance for NLP standardization (Participant 10) due to Punjabi's unstandardized nature. The analyzed data demonstrates the clear necessity for dialect-empowered AI systems while requiring additional Punjabi language vocabulary along with AI recognition training and uniform data standardization for Punjabi NLP to achieve a correct linguistic and cultural result in language revitalization efforts.

Table 3. Policymakers Data analysis

Participant Number	**Participant Role**	**Key Theme**	**Key Insights**	**Recommendations**
11	Policymaker	Educational Policy	Punjabi lacks formal integration in Pakistan's education system, limiting digital resource development.	Mandate Punjabi language inclusion in AI research and educational technology programs.
12	Policymaker	Digital Language Justice	Lack of AI-driven initiatives for Punjabi contributes to its digital exclusion.	Promote AI policies that ensure digital inclusivity for Punjabi and other underrepresented languages.
13	Policymaker	Public Awareness	Limited public awareness about AI's role in language revitalization restricts policy-driven efforts.	Develop national awareness campaigns on AI-driven language preservation.
14	Policymaker	Funding	Funding for language-focused AI projects remains	Allocate government funds to AI projects

			negligible compared to other tech domains.	focused on low-resource languages like Punjabi.
15	Policymaker	Government Collaboration	No structured collaboration exists between government bodies and AI researchers for Punjabi language preservation.	Foster collaboration between policymakers, linguists, and AI developers for structured language revitalization.

Research findings about policymakers identify major policy hollows and system limitations in Punjabi language revitalization through AI-driven solutions that match Fishman's (1991) Reversing Language Shift (RLS) Model as well as Technological Determinism Theory (McLuhan 1964) and Digital Linguistic Justice (Bender et al., 2021). The Pakistan formal education system ignores Punjabi which prevents the creation of AI-based digital resources making it harder to integrate AI into educational programs according to Participant 11. Participant 12 points to digital language justice deficiency which results in further marginalization of Punjabi in AI-powered applications because low-resource languages face digital exclusion through policy and technological choices.

Insufficient community understanding (Participant 13) stops AI policy-based implementations from taking effect yet surveys advocate for AI education drives targeting the Punjabi language group. Funding limitations for Punjabi AI innovation projects (Participant 14) resonate with the worries expressed by experts in both fields who encourage consistent financial backing for NLP tool development and speech recognition and machine translation system creation. The conclusion between government agencies and AI experts and linguists remains unstructured thereby creating an incomplete link between policy evolution and technological development (Participant 15). This confirms the questionnaire's request for co-operative research between different disciplines to successfully implement Punjabi language preservation through AI solutions. The present analysis demonstrates a compelling requirement for educational institutions, policy revisions, allocated funding and AI-based public involvement and collaborative governance to maintain AI as a sustainable force for Punjabi language preservation.

Table 4. Community Activists Data analysis

Participant Number	Participant Role	Key Theme	Key Insights	Recommendations
16	Community Activist	Cultural Identity	Punjabi is being replaced by Urdu and English due to socio-political pressures, affecting language sustainability.	Encourage AI-assisted community language learning initiatives and online Punjabi learning platforms.
17	Community Activist	Grassroots Movements	Local advocacy movements for Punjabi language revival lack digital tools and technological support.	Develop AI-driven cultural content (music, poetry, storytelling) to engage younger audiences.
18	Community Activist	Digital Engagement	AI-based Punjabi language tools are underutilized due to limited awareness and accessibility.	Create awareness about AI-powered Punjabi translation and speech tools for digital empowerment.
19	Community Activist	Community Resources	Community resources for Punjabi learning are not integrated with modern AI-driven platforms.	Support grassroots movements in creating AI-integrated Punjabi linguistic archives.
20	Community Activist	Language Transmission	Younger generations lack motivation to learn Punjabi due to limited career benefits in formal domains.	Advocate for AI tools that make Punjabi learning fun, interactive, and career-relevant.

The analysis of community activists reveals linguistic obstacles blocking Punjabi language survival while AI holds promise to counteract these negative developments by following Fishman's (1991) Reversing Language Shift (RLS) Model and Technological Determinism Theory by McLuhan (1964) as well as Digital Linguistic Justice by Bender et al (2021). The disappearance of cultural identity (Participant 16) occurs because Urdu and English have gained dominance and follows broader patterns of language shift that also worry respondents about Punjabi's intergenerational survival. The lack of digital tools inside grassroots movements (according to Participant 17) creates obstacles for Punjabi revival and demonstrates the importance of AI-based cultural content (music, poetry, storytelling) to attract youth members as theoretical perspectives on digital empowerment indicate for language sustainability. User Engagement with digital Punjabi language tools remains poor because most people are unaware of these tools or cannot access them as shown by Participant 18 who represents this issue.

This finding confirms questionnaire results regarding the necessity for general public advertising about AI-driven linguistic tools. Participant 19 illustrates a crucial space where artificial intelligence has not been integrated into community resources thereby exposing a distinctive divide between classic language advocacy methods and contemporary technological approaches while supporting the main points from the questionnaire about implementing AI technology in Punjabi educational initiatives. AI solutions must provide interesting and profession-boosting Punjabi learning activities since young people show diminished motivation (Participant 20) to study the language because of its limited career value. The research demonstrates that AI-powered community-based projects need to unite with knowledge campaigns and digital education platforms and AI-aided linguistic repositories to secure Punjabi language survival in digital environments.

Table 5. Tech Developers Data Analysis

Participant Number	Participant Role	Key Theme	Key Insights	Recommendations
21	Tech Developer	AI Infrastructure	Limited AI infrastructure in Pakistan slows down NLP advancements for Punjabi.	Invest in AI infrastructure for language processing, including cloud-based NLP training environments.
22	Tech Developer	Data Scarcity	Scarcity of high-quality training datasets affects AI model performance.	Develop structured Punjabi linguistic datasets to enhance AI learning capabilities.
23	Tech Developer	Speech-to-Text Models	Existing speech-to-text models fail to capture Punjabi phonetic nuances accurately.	Refine AI models for Punjabi speech-to-text accuracy through phonetic and dialectal adjustments.
24	Tech Developer	Model Adaptation	Current AI models need adaptation for low-resource languages like Punjabi.	Create AI systems that adapt existing NLP architectures for Punjabi.
25	Tech Developer	Open-Source Development	Few open-source Punjabi AI projects exist, limiting collaborative advancements in language technology.	Encourage open-source contributions to build a collaborative Punjabi AI ecosystem.

The research of tech developers identifies technical challenges and AI-based solutions for Punjabi language revitalization while following the interpretations of Technology Determinism Theory (McLuhan, 1964), Fishman's (1991) Reversing Language Shift (RLS) Model and Digital Linguistic Justice (Bender et al., 2021). The insufficient AI infrastructure throughout Pakistan detected by Participant 21 hinders Punjabi NLP innovation while this finding validates the questionnaire interviewees' points regarding specially designated AI infrastructure needed for linguistic processing. High-quality Punjabi datasets remain scarce (Participant 22) which leads to reduced AI model effectiveness since AI algorithms need extensive annotated data to work efficiently.

This observation strengthens the linguists' worry about limited datasets presented in the questionnaire. Research conducted by linguists about AI struggling with Punjabi phonetics and language dialects finds support in Participant 23 describing poor speech-to-text model accuracy. Models need to be adapted because contemporary Natural Language Processing systems operate

exclusively on resources-laden languages (Participant 24) thus confirming concerns expressed in the questionnaire about AI system biases toward under-resourced languages such as Punjabi. The unavailability of open-source Punjabi AI projects (as reported by Participant 25) prevents collaborative work in language technology thus supporting community activists who advocate for open-access AI projects to preserve the Punjabi language. The analysis demonstrates the critical requirement for AI infrastructure development and data expansion as well as speech recognition system enhancement together with the creation of adaptive NLP models within open-source AI ecosystems to keep Punjabi language sustainable online while preventing new language exclusion in AI development.

The research shows AI demonstrates substantial power for revitalizing Punjabi language but the key obstacles throughout its path refer to data accessibility challenges alongside technology restrictions and absent policies and under engaged participants and development limits of infrastructure. Technology research experts and developers through questionnaire responses identified data limitations together with missing Punjabi-specific Natural Language Processing models as principal challenges that strengthen the necessity to create standardized linguistic databases and adapt NLP models to dialects and enhance speech recognition. Components like complex morphology and code-switching behavior and lexical gaps in Punjabi language demand the creation of AI-powered tools that can represent multilingual diversity properly. The lack of government AI support together with funding deficits and insufficient public awareness were identified by policymakers as factors that require educational digital language justice programs and government backing for such initiatives. Community activists demonstrated how Punjabi faced decline due to social and political factors so they advocated for AI-based local participation and cultural content development strategies that engaged the youth population. Tech developers expressed concerns about deficient infrastructure together with model adaptation difficulties and open-source AI absence and therefore demanded financial investments for AI resource development along with collaborative initiatives. The study establishes AI as a vital tool for reviving the Punjabi language yet stresses that sustained impact requires joining different groups of stakeholders who combine technological progress with proper policy guidelines and active community participation along with responsible AI technology growth.

DISCUSSION

The problems with AI-driven Punjabi language renewal emerge from historical mistreatment together with linguistic dominance and technological disconnect which become worse thanks to insufficient government policies and system-based AI processing weaknesses and insufficient community-led digital advancement programs. Currently the most pressing challenge emerges from Pakistan's official bureaucracy which fails to give formal legal status to Punjabi across educational settings and also within government institutions and digital policy development spheres. The educational curriculum in Pakistan does not teach Punjabi whereas policymakers (Participant 11) state that this language lacks structured development needed for AI adaptation. By being officially excluded Punjabi registers show diminished visibility while scheduled development of AI-driven linguistic resources is blocked since NLP applications need standardized language corpora for effective operation. The lack of government-funded AI projects for Punjabi language contributes to digital marginalization which leads to Urdu and English languages dominating artificial intelligence research and technological development (Participant 15). The ability to preserve and revitalize the Punjabi language through AI technology is limited when the state does not provide policies funding and infrastructure which supports such efforts. External actors in Punjabi digital development leads to disorganization and inconsistent behavior which obstructs the ability to establish sustainable growth and maximize scalability in Punjabi digital environments.

The ethical problems in Punjabi language preservation through AI depend on issues of algorithmic bias and linguistic standardization and digital control which reduce the efficiency of AI systems to sustain the language. AI studies specialists (Participant 3) have identified how prevailing AI models exclusively work with English and Mandarin languages which creates an architectural exclusion for Punjabi from mainstream NLP research developments. Digital linguistic justice faces critical issues (Bender et al., 2021) because AI systems provide better services to languages that have abundant training datasets that Punjabi does not have access to thus it becomes algorithmically marginalized. The absence of Punjabi data specific datasets discussed by Participant 1 significantly hampers AI's capability to understand Punjabi language content leading to unsatisfactory results during

automated linguistic processing. Participant 10 among the linguists stated that Punjabi experiences major challenges in developing standardized digital corpus which creates barriers for AI models to correctly process Punjabi dialects and complex grammar patterns along with code-switching usages. Digital colonialism exists according to Tech developers (Participant 25) because AI technologies from Global North countries impose their control over Punjabi digital representation to exclude native speakers and Punjabi linguists, developers from shaping AI linguistic tools. Running AI development from top-level sources prevents native expertise integration thus it deepens Punjabi's digital isolation and becomes an obstacle for its success in AI networks. Equitable sustainable digital integration of Punjabi requires the involvement of multiple stakeholders who should work on implementing policy changes and driving AI development through local communities while standardizing the language and establishing open AI collaborations.

The potential of Artificial intelligence technologies including Natural Language Processing (NLP), speech recognition and machine translation to revitalize the Punjabi language remains limited because of technology-related obstacles together with socio-political and linguistic challenges. The underdeveloped application of AI for Punjabi language preservation is limited by the shortage of available data and standardization issues and institutional lack of focus. AI researchers together with linguists along with policymakers and tech developers and community activists present a network of possibilities and obstacles when applying AI for Punjabi language digital viability. AI technology creates three main benefits for Punjabi digital accessibility by building automatic transcription services alongside digital learning tools and AI content generation assistance to expand Punjabi online visibility. The combination of speech recognition and NLP models enables the documentation process while allowing transcription and digitalization of Punjabi so it persists across communication and educational and media platforms.

The fourth AI researcher explained that AI models function similarly to successful projects in Te Reo Māori and Welsh by learning Punjabi dialects in order to advance community use of indigenous languages, although such models require specialized training. Machine translation platforms created with AI neural network technology provide linguistic connection between Punjabi and worldwide dominant languages to increase multilingual access. According to Linguists (Participant 8) wider linguistic databases would boost machine translation precision so Punjabi speakers could easily join international digital connection networks through their language. Participant 17 explained that digital storytelling along with poetry and musical entertainment enhanced through AI produces content to rekindle Punjabi interest among younger audiences who also utilize technology platforms for its transmission to subsequent generations.

AI's efforts to save the Punjabi language face serious challenges resulting from insufficient language data along with complex dialects and biases in machines and lack of government-led AI programs. AI researchers stressed that Punjabi language lacks formal annotated linguistic resources because this hinders development of AI systems that properly understand Punjabi syntax, phonetics and semantics (Participant 1). The AI models operated by Tech developers (Participant 22) need extensive linguistic data to function properly but Punjabi lacks adequate datasets in comparison to English and Urdu along with Hindi which have substantial government-sponsored NLP resources. The complex nature of Punjabi language together with its various dialects creates major obstacles during AI adaptation processes. AI faces substantial problems while processing Punjabi syntax as well as morphology and code-switching patterns due to which it creates translation mistakes and weak NLP performance and generates subpar AI text according to Linguists (Participant 6). The speech recognition models of Participant 23 feature a fundamental flaw that causes them to fail when identifying different Punjabi dialects which leads to reduced relatability for multifaceted speaker communities. Dialect-sensitive training data sets are necessary to prevent Punjabi from becoming overly simplistic or being misrepresented or combining dialects since such approaches can lead to native speaker rejection instead of encouraging language adoption.

The persistent bias within AI systems continues to discriminate against underprivileged languages such as Punjabi when exclusive attention is focused on dominant global languages together with their corresponding languages. The AI researchers (Participant 3) highlighted how current AI technologies prefer English together with Mandarin and other resource-rich languages to neglect proper representation of Punjabi in worldwide NLP developments. The concept of digital linguistic justice described by Bender et al. (2021) states that AI language tools without inclusive policies

would increase language inequalities instead of providing solutions to these digital divides which threaten Punjabi languages and other minority languages. Public AI initiatives for Punjabi do not exist and this fact results in weak institutional support for Punjabi digital representation because private sector funding and community initiatives now maintain Punjabi AI solutions without institutional infrastructure development (Participant 12).

Social, political and decreasing public interest creates additional barriers for Artificial Intelligence to fully flourish. Lack of sufficient public funding for Punjabi AI-driven projects restrains researchers in the development of NLP and speech recognition technology and AI-based digital applications (Policy-maker Participant 14). Changing incentives for career success has led to reduced motivated learning of Punjabi among young people according to activists in the community (Participant 20) who noted this will hinder the adoption of AI-based solutions. A combination of multi-stakeholder involvement with AI applications in education and employment and digital spaces remains essential to stop the decline of Punjabi language because stand-alone technological advances are inadequate. Such intervention demands strategic policies and AI research programs along with active community support systems for Punjabi language protection.

A successful revival of Punjabi language through artificial intelligence demands an approach which unites stakeholders from various disciplines that include technological developers alongside linguists and policymakers and local language communities. Technical developers stressed the need for an open-source AI platform because this would enable linguistic researchers and AI scientists as well as Punjabi community members to work together for better Punjabi NLP model development (Participant 25). A mandatory requirement for AI models is their ability to identify Punjabi dialectical diversity because this prevents linguistic basification while ensuring correct representation of Punjabi grammatical structures and phonetic variations with its regional dialects (Participant 24). The policymakers (Participant 15) emphasize the importance of receiving state-funded AI research grants alongside institutional partnerships between policymakers and AI developers to enable government-backed funding for platforms preserving languages through AI technology.

Marginalization of Punjabi language by the state prevents AI-based language revitalization models from developing at scale because private-sector and grassroots projects fail to match dominant global language capacities. Public engagement and AI literacy education programs (Participant 13) should become essential for building Punjabi speaker involvement in AI dataset development and testing and language research thus enabling AI to emerge as a community-formed technology instead of an implementing authority tool. AI functions properly as a support system alongside human efforts to maintain Punjabi while educational changes happen alongside media programming and grassroots support against language decline in digital environments.

The extensive capability of AI-driven technologies to revitalize the Punjabi language remains constrained because of four main structural issues involving inadequate data collection and difficulty with dialects and AI system biases and minimal public sector support. Researchers working in AI (Participant 1) and tech developers (Participant 22) agree that Punjabi lacks organized linguistic datasets with annotations because this makes training AI models which properly process Punjabi syntax, phonetics and semantic elements very difficult. Punjabi stands apart from popular languages like English along with Mandarin and Hindi since it receives minimal AI-based linguistic study along with sparse linguistic resources and insufficient technical infrastructure which intensifies digital isolation. AI models prove difficult to distinguish Punjabi dialects according to linguists and speech recognition specialists as this inability results in poor NLP performance and unreliable speech-to-text conversion and unsatisfactory machine translation.

If AI technology does not use dialect-specific models along with standardized linguistic resources then Punjabi faces three possible outcomes: systematic misrepresentation through integration or single standardization or digital exclusion from all language initiatives. Research on digital linguistic justice by Bender et al. (2021) states that artificial intelligence language technology more often benefits dominant languages than language minorities thereby establishing linguistic inequality in the place of addressing it. Heuristic system development for Punjabi receives almost no support from government funding which blocks the chance for developing meaningful Natural Language Processing solutions beyond restricted dataset growth (Participant 14 testified as a policymaker).

Community activists suggest that young people do not show interest in learning Punjabi because there are no significant career advantages related to the language thus making AI-based solutions potentially useless for revitalization (Participant 20). The efforts to revitalize the Punjabi language through AI must be incorporated into complete linguistic policies along with educational frameworks and socio-cultural initiatives that dedicate themselves to linguistic equity and ethical AI development and sustained community participation.

Findings of the Study

The research shows that AI technology applications present revitalization potential for Punjabi language yet they function poorly due to deficient linguistic resources and complex language structure together with insufficient policies and exclusion from digital spaces. Insufficient structured linguistic corpora for Punjabi has resulted in major operational impediments for NLP systems together with speech recognition and machine translation tools which then make AI processes incapable of handling Punjabi syntax structures and phonetic patterns alongside dialectical changes accurately. The present AI models systematically select high-resource languages over Punjabi therefore preventing its access to worldwide NLP innovations because the language does not have adequate training data and faces algorithmic bias. Linguistics professionals pointed out that Punjabi language contains intricate morphological features and multifaceted inflections along with code-switching patterns that AI systems find challenging to process unless standardization of datasets and dialectal adjustments occur.

Government policies have shown no interest in AI-based Punjabi language initiatives resulting in language preservation depending mostly on grassroots movements together with private-sector innovation instead of institutional backing. The decline of Punjabi happened due to socio-political influences and community activists agreed that AI tools alone are insufficient to stop language shift unless they combine educational integration and cultural engagement and public awareness initiatives. The lack of infrastructure support and the dearth of open-source AI tools and collaborative platforms according to tech developers presents challenges for digital sustainability of Punjabi language. The research demonstrates that AI requires support from an integrated solution that combines dialect-specific AI model creation and government policy frameworks together with structured data development as well as community involvement initiatives for Punjabi language survival during the AI era.

The research output matches prior studies examining AI-based low-resource language maintenance which proves that AI technologies work ineffectively to stop language loss alongside minimal structured linguistic databases as well as insufficient institutional aid and community grassroots involvement. According to Jones and Hammarström (2021), Bird (2020) along with AI Researchers (Participant 1) indicate how NLP models using AI need substantial linguistic datasets of high quality which are currently absent in the digitized corpora of Punjabi. The digital linguistic justice studies conducted by Bender et al. (2021) demonstrate that AI systems demonstrate a built-in preference toward high-resource languages thus excluding Punjabi from global AI evolution according to technical developers (Participant 22).

Findings from research about Māori and Welsh AI revitalization activities (Keegan, Mato, & Ruru, 2015) prove how policy support and community engagement lead to effective AI approaches and match the worries voiced by officials (Participant 14) about Punjabi AI project absence. The participants who were linguists (Participant 9) in this research confirmed established findings (Hellwig, 2018) about AI's problems with analyzing complex language features including code-switching alongside phonetic variations which also appeared in AI technology used for African and South Asian mother tongues. Research by Blodgett et al. (2020) shows how AI models fail to accommodate dialectal diversity which matches the observations of Punjabi speech-to-text errors made by AI researchers (Participant 4) in this study. The study confirms the consensus between AI and linguistics researchers about how digital disparities persist in Punjabi because standardization is absent alongside inadequate training datasets and ethical AI development and state involvement.

The conclusions of this research strengthen Fishman's (1991) Reversing Language Shift (RLS) Model and Technological Determinism Theory (McLuhan, 1964) as well as Digital Linguistic Justice (Bender et al., 2021) because they demonstrate AI requires institutional support alongside technological advancements and community leadership for Punjabi revitalization. The research results support

Fishman's RLS Model since this study found policy support and intergenerational transmission to be essential for revitalization according to policymaker Participant 14 who noted that absent government support Punjabi revitalization depends on community-driven action. The study presentation shows how limited AI infrastructure and lack of Punjabi datasets influence the inability of this language to integrate into digital platforms. Implementation and technical success of AI depend on institutional support which this research demonstrates through interviews with AI research specialists (Participant 1) and software developers (Participant 22). The Digital Linguistic Justice framework (Bender et al., 2021) states that AI systems prefer dominant languages above other languages which leads to exclusion from digital development according to AI researcher Participant 3. He reported that Punjabi faces marginalization in global NLP models because of algorithmic bias and data scarcity. The linguists described the struggle of AI with Punjabi while emphasizing both the complexity of its syntax and language dialects thus advocating for proper ethical AI advances (Participant 6). This research confirms that Punjabi digital marginalization cannot be solved through AI technology without funding and emerging datasets alongside community inclusion according to the theoretical frameworks adopted for this study.

Implications of the Study

The study results generate important consequences regarding future work for AI developers together with linguists and both policymakers and language advocates. AI developers and researchers should make Punjabi NLP model adaptations their highest priority through the improvement of speech recognition systems and coding of expansion datasets that should integrate with code-switching and regional speech variations. Political leadership should unite AI technology with language planning while creating state-funded programs for Punjabi digital inclusion and allocating AI research funds as well as institutional support. The standardization of Punjabi linguistic resources by researchers and linguists should become their priority while they work on resolving orthographic conflicts between Shahmukhi and Gurmukhi and grammatical structure issues to enhance Punjabi AI compatibility. A comprehensive plan should be established to make AI-driven language tools freely available to the public through community awareness programs which promote their cultural relevance among people who speak the language.

Limitations of the Study

The study delivers extensive evaluation although researchers need to recognize several constraints. The research bases its findings mainly on qualitative data collected from AI researchers together with linguists' policymakers' community activists and tech developers instead of performing technical AI model tests that could confirm the NLP limitations of Punjabi language processing by AI systems. The research fails to explain how the script-based differences between Shahmukhi in Pakistan and Gurmukhi in India affect the development of AI systems across different geographical regions. AI adoption tends to differ differently in urban and rural areas as well as local/global settings due to the major differences in digital presence of Punjabi between Pakistan and the diaspora. Future research must include AI performance evaluation measures while developing applications across different Punjabi-speaking areas and analyzing AI performance in contrast to other minority languages to enhance these studies.

CONCLUSION

Research analyzed AI technology participation in Punjabi language rebirth by investigating advantages together with disadvantages. The data deficiency alongside linguistic barriers and AI operational problems combined with institutional backing shortages limits the impact that NLP, speech recognition and machine translation systems can make toward Punjabi digital sustainability. The absence of standardized linguistic corpora in Punjabi results in poor performance of AI models according to developers and researchers of artificial intelligence. Linguists pointed out that Punjabi language features complicated morphology alongside various inflections combined with different dialects which poses problems for AI-based processing unless developers create specific AI models for each dialect. The policy makers highlighted that the national government does not support AI-enabled Punjabi protection initiatives because such programs lack institutional backing. Socio-political pressures coupled with public disengagement makes it doubtful that AI-based tools would find wide adoption even if they become available because community support and career incentives for their implementation are lacking. The research confirms AI stands alone as an insufficient

solution for Punjabi language revitalization although it remains a potent tool that requires combination with policy changes and research about language alongside community initiatives and ethical AI development for lasting digital Punjabi inclusion.

The presented investigation identifies multiple important research topics for future investigation. The development of extensive Punjabi linguistic datasets across all dialects remains a first research task for AI experts to enhance NLP and speech recognition and machine translation systems accuracy levels. Governments along with institutions need to boost their collaboration efforts with AI developers who need financial and policy backing to drive Punjabi revitalization through artificial intelligence applications. Public engagement together with open-source AI initiatives must be encouraged because they allow Punjabi speakers to actively participate in language preservation projects through AI technology. Future research needs to establish ethical guidelines for AI development so digital linguistic justice becomes possible thereby ensuring Punjabi integrates fully as a digital language.

AI emerges as a possible revolutionary language revitalization tool for Punjabi when it receives ethical development through partnerships with linguists and policymakers and the community. AI systems will fail to become an accessible solution for Punjabi language preservation unless available linguistic data is structured while there is dialect-sensitive adaptation along with government funding and strong public engagement. The success of AI as a sustainability bridge for Punjabi language depends on collaborative involvement between stakeholders to prevent it from blocking digital access for Punjabi speakers across the globe.

REFERENCES

Abdullah, H. (2023, January 18). UBC's 'Punjabi in BC' project preserves the efforts of those who have made the revival of the Punjabi language possible. *5X Fest*. https://www.5xfest.com/5xpress/ubcs-punjabi-in-bc-oral-history-project-preserves-the-efforts-of-people-who-have-made-the-revival-of-the-punjabi-language-possible

Ayres, A. (2009). *Speaking Like a State: Language and Nationalism in Pakistan*. Cambridge University Press.

Aziz, E. A. (2024, September 5). Endang Aminudin Aziz. *TIME*. https://time.com/7012839/endang-aminudin-aziz/

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.

Bird, S. (2020). Decolonising speech and language technology. *Proceedings of the 28th International Conference on Computational Linguistics*, 3504–3519.

Blodgett, S., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5481–5497.

Butt, M., Ashraf, M. J., Sanjarani, K. M., Khan, M. A., & Arshaad, T. (2024). The politics of language and the fall of Dhaka: A discourse historical analysis. *Migration Letters, 21*(S14), 1065–1084.

Crystal, D. (2000). *Language Death*. Cambridge University Press.

Feenberg, A. (2010). *Between Reason and Experience: Essays in Technology and Modernity*. MIT Press.

FirstVoices. (2023). *FirstVoices*. https://www.firstvoices.com

Fishman, J. A. (1991). *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Multilingual Matters.

Hellwig, O. (2018). Challenges in developing NLP tools for under-resourced languages: The case of Sanskrit. *Language Resources and Evaluation*, *52*(1), 71-92.

Hovy, D., & Spruit, S. L. (2016). The social impact of natural language processing. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 591–598.

Hussain, F., Khan, M. A., & Khan, H. (2024). Issue 1 Punjabi Language Perspectives: Marginalization and Advocacy in Pakistan. *ResearchGate*. https://www.researchgate.net/publication/384723699_Issue_1_Punjabi_Language_Perspectives_A_Advocacy_in_Pakistan

Hussain, F., Razzaq, S., Butt, M., Sarwar, T., & Khan, M. A. (2024). The evolution of language politics: A CDA study of Urdu and Punjabi in Pakistan. *History of Medicine, 10*(2), 956–980.

Jones, B., & Hammarström, H. (2021). NLP for language revitalization: Opportunities and challenges. *Journal of Linguistic Technology*, *5*(2), 44-62.

Jones, P.-L. (2024, September 5). Peter-Lucas Jones. *TIME*. https://time.com/7012841/peter-lucas-jones/

Keegan, T. T., Mato, P., & Ruru, S. (2015). Using digital technology to support Māori language revitalization. *International Journal of Indigenous Language Research*, *7*(3), 56-74.

Mansoor, S. (2005). Language planning in higher education: A case study of Pakistan. *Oxford University Press*.

McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. MIT Press.

O'Regan, H. (2022). AI and indigenous language survival: A critical perspective. *Language and Society Review*, *10*(4), 27-39.

Postman, N. (1992). *Technopoly: The Surrender of Culture to Technology*. Knopf.

Rahman, T. (2011). *Language Policy, Identity and Religion: Aspects of the Civilization of the Muslims of Pakistan and North India*. Oxford University Press.

Shackle, C. (2003). Punjab and the rest: Language and identity in Pakistan. *Modern Asian Studies*, *37*(3), 715-740.

Zulfiqar, I., Ashraf, M. J., Khan, M. A., Abbas, Z., & Saeed, M. U. (2024). The politics of language in rural Punjab: Exploring social stratification, linguistic capital and power dynamics through critical discourse analysis. *Remittances Review,* *9*(S3), 1316–1351. https://doi.org/10.33282/rr.vx9i2.63