Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

AI-Powered Multilingual Voice Banking Assistant

Khushi Agrawal
Department of Artificial Intelligence & Data Science, Thakur College of Engineering and Technology, Mumbai, India

Dhruv Bhandari
Department of Artificial Intelligence & Data Science, Thakur College of Engineering and Technology Mumbai, India

Yash Berde
Department of Artificial Intelligence & Data Science, Thakur College of Engineering and Technology Mumbai, India

Prof. Swati Joshi
Department of Artificial Intelligence & Data Science, Thakur College of Engineering and Technology Mumbai, India

Abstract- ***The development of digital banking has enabled access to financial services, but some issues related to digital literacy, language problems, and complicated interfaces have still been encountered by specific users, especially older people and users without much experience in using digital services. This paper provides an AI-Powered Multilingual Voice Banking Assistant aimed at facilitating banking operations by using voice commands. The suggested system includes speech-to-text conversion, multilingual processing, intent recognition based on keywords, backend processing, interaction with a database, and text-to-speech feedback that allow communicating in English, Hindi, and Marathi. With the help of the prototype, users can get access to their accounts, view their balances, obtain transaction history, perform money transfers, pay money to saved beneficiaries, and obtain bank statements via voice interaction. A synthetic dataset for banking is used for implementing the application and conducting functional tests. Evaluation of the suggested solution includes checking voice command recognition, intent recognition, performing banking operations, and response time in three languages. It has been shown that most of the tested voice commands have been successfully processed. The speech input can be influenced by the noise in the environment, pronunciation, accents, and speech quality; however, it is clear that the system proves the potential of voice-assisted technologies to make crucial banking services more accessible. With the help of multilingual interaction combined with typical banking functions, the suggested system proposes an alternative interface for those users who might have problems using traditional banking applications.***

Keywords: Multilingual Speech Processing, Voice-Based Banking Systems, Automatic Speech Recognition (ASR), Natural Language Processing (NLP), Intent Classification, Artificial Intelligence (AI), Voice User Interface (VUI), Conversational AI Systems, Human–Computer Interaction (HCI), Digital Financial Inclusion

I. INTRODUCTION

With the emergence of digital banking, the process of delivering banking services has been revolutionized, which has improved the speed and availability of banking. However, there are still issues faced by rural people and senior citizens when using digital banking services because of digital illiteracy, linguistic differences, lack of knowledge regarding smartphones and complexities associated with the applications [1], [2]. Rural users' issues in adopting and using digital banking services in the state of Maharashtra have been reported by researchers as well [1].

Traditional mobile banking apps usually involve reading instructions, entering information, and navigating through various menus and screens. It becomes a challenge for elderly people and those who have less knowledge about digital technology. According to recent research on the usage of FinTech by senior citizens, digital literacy is very important for the acceptance and use of mobile banking technology [2]. This issue causes confusion, mistakes, loss of confidence, and dependence on family members and intermediaries for basic financial transactions.

The progress in AI, ASR, and NLP has given rise to new possibilities in overcoming some of these limitations via voice interactions. Voice-based interface allows the user to interact with the system using spoken language, which does not fully depend on reading, writing, and graphical navigation. In previous research, voice-based banking for rural and semi-literate users has been considered, showcasing the possibility of providing financial services via speech and local languages [3].

But developing a reliable voice interface for Indian users is even more complicated due to linguistic variety in India, regional accent and pronunciation problems, and common practice of speaking several languages at once. Multilingual Indian ASR research has shown the problems of providing ASR support for several low-resource Indic languages and code-switching [4]. More specifically, recently conducted research on Hindi–Marathi code-switching ASR has proven that linguistic variety, phonetic differences, and lack of speech resources are still major problems in this field [5]. The availability of multilingual speech resources like SPRING-INX clearly shows the necessity of developing speech technology for Indian languages, including Hindi and Marathi [6].

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

II. RELATED WORK

Conversational AI, speech recognition, and digital banking research has looked at ways of achieving better access to banking services through various techniques. Current literature has considered such issues as banking assistants, chatbots for customer service, speech processing in multiple languages, and voice interface usability. All of them form good bases for voice-enabled banking software development.

A. Conversational Banking Systems

Bhatia and Khetarpaul [7] have proposed the development of an AI-powered banking assistant which is proficient in English, Hindi, and Hinglish. In their system, the process of language detection, AI response generation, and voice input/output are used to answer questions about the banking industry. The success of their system has been measured by testing the application among 50 participants over 100 sessions with an 87% success rate. This study shows the capability of multilingual conversation systems in helping out banking-related tasks. But, their work was based on conversational assistance and bank information, whereas our research emphasizes on database transactions through conversations.

In addition, Saranya et al. [9] analyzed the application of an AI-based voice assistant for banks which integrates speech recognition, NLP, banking operations, and security features. This is related to our task of combining voice interaction and banking services. In comparison with those approaches, there is one key element of system design which we should pay attention to: a bank's voice assistant should be able not only to receive the user's voice command but also to understand it and to perform the necessary banking operation.

B. Conversational Agents in Financial Services

In addition to chatbots and voice agents in the context of business-customer interaction, Bălan [8] analyzed the opportunities offered by them to enhance communication and customer care, as well as the obstacles that may occur when using them because of the issues related to trust, privacy, and acceptance of users. Eustaquio-Jiménez et al. [13] carried out a systematic review of the usage of chatbots in financial organizations, focusing on the implementation of chatbots in customer support and solution of various issues. In the same way, Graham et al. [14] explored the usage of chatbots in the sphere of banking and finance.

In summary, these works show the importance of conversation agents for interacting with customers and in the domain of finance. Nevertheless, in case of general chatbots and in case of banking operations, there are different implementation challenges. It is necessary that the chatbot designed for banking tasks will link users' requests to particular back-end operations and provide answers based on the corresponding application information. Furthermore, the interaction mechanism of the chatbot should be considered independently from the correctness of the banking operations themselves. The current research is concerned with the above integration in the web-browser prototype.

C. Multilingual Speech Recognition and Voice-Interface Usability

The importance of multi-lingual speech recognition becomes particularly significant when it comes to speech recognition systems meant for communities of varied languages. The issue has been examined by Diwan et al. [4], who have researched multilingual and code switching automatic speech recognition (ASR) in under-resourced languages of India. In the course of this study, the problems that occur because of the difference in the linguistic, acoustic, and speech resource properties become clear. Palivela et al. [5] have also performed research on Hindi-Marathi code-switching ASR, highlighting the technical issues related to speech recognition in several languages.

These researches mostly deal with speech recognition and speech resources, but not the identification of intents of banking request. This difference becomes crucial for the design and assessment of voice banking assistant. ASR performs conversion from the speech form to the text form, and intent identification identifies an operation that the text corresponds to. In the described prototype, speech recognition is done using speech technologies implemented in a web browser, and intent identification is performed on the basis of the keywords and rules. No training of ASR model or machine learning classifier of intents is done in this prototype.

Usability is another factor that affects the development of voice-based applications. Design guidelines for speech-recognition interfaces were evaluated by Jones et al. [10], and Wolters et al. [11] studied the interaction of older people with spoken-dialogue systems. Human factors in relation to voice-user interfaces and older adults were assessed by Stigall and Caine [12]. Such research provides useful background in designing understandable prompts and easy-to-use voice interaction. In case of developing a banking assistant that will take into account various user experiences with digital technologies, usability is one of the key aspects to consider. However, this is what should be tested separately.

D. Research Gap and Contribution

The literature surveyed encompasses different perspectives related to voice banking, with conversational banking systems showing how the technologies are implemented in finance [7], [9], conversational-agent studies focusing on customer service applications and implementation issues in general [8], [13]–[15], and multilingual ASR and voice-interface studies focusing on speech recognition and interaction issues [4]–[6], [10]–[12]. This body of literature serves as a basis for the current paper; however, the conclusions from the studies cannot be used to suggest that all banking assistants lack multilingual capabilities and certain banking functions.

On the other hand, this research emphasizes the implementation and evaluation of a particular set of functionalities within a particular browser-based prototype. The prototype supports English, Hindi, and Marathi languages. It is capable of speech-to-text conversion, keyword and rules-based intent recognition, banking functionalities using database, and text-to-speech feedback.

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

The functionalities include inquiry about account balance, transaction history retrieval, funds transfer, payment to five saved beneficiaries of the primary user, and downloadable statements.

The importance of this research is that it captures the integration and evaluates the system via functional testing and performance timing in the three supported languages. This is the first level of evaluation of end-to-end interaction latency and does not itself prove the ASR accuracy, intent identification accuracy, usability for older or rural populations, or even security in production. All of these continue to be areas worthy of future study.

III.DATASET PREPARATION AND PREPROCESSING

A. Dataset Overview

The AI-powered Multi-Language Voice Bank Assistant makes use of an artificial dataset that was designed for implementation and testing purposes. The dataset includes three files: users, transactions, and saved contacts. Each file represents information on user accounts, transactions, and benefactors, respectively. A brief description of the dataset format and purpose of each file can be found in Table I.

Table 01:Synthetic datasets used in proposed system

Dataset	Records	Role in preparation
Users	300	User profiles and account information
Transactions	10,000	Transaction histories of primary users
Saved contacts	250	Beneficiary relationships

B. Source data and Attributes

In the users database, there are 300 unique user records that consist of 50 primary bank users and 250 users having saved beneficiary contacts. The database has user id, contact details, and account information needed to model users.

The data set for transactions is made up of 10,000 records, which are divided into 50 main banking users that do 200 transactions each. Every record contains many unique attributes about the transaction, such as the transaction ID, user ID, time and date, transaction type, amount of transaction, beginning balance, ending balance, payee, and mode of transaction. Transaction types define whether the transaction is done by the credit and debit process. However, mode of transactions shows different types of banking processes, such as UPI, debit card, NEFT, ATM, net banking, bank transfer, cash, bank credit, and UPI payees.

Contacts dataset contains 250 records where there are five contacts for each of the main user from banking database. In every record there is identifier of both the main user and the contact user and also the contact information such as the saved nick name. The dataset contains relationships required to link the beneficiary contacts with their corresponding main users.

Table 02:Structure and Attributes of the Synthetic Banking Dataset

Dataset	Key Attributes	Purpose
Users	User ID, name, contact details, account details	Stores user profiles and account related information
Transactions	Transaction ID, User ID, date, time, transaction type, amount, previous balance, balance after transaction, merchant/payee, transaction mode	Records banking activities and supports transaction history retrieval
Saved Contacts	Primary User ID, Contact User ID, saved nickname, contact details	Represents beneficiary relationships and supports payments to saved contacts

C. Data Integration and Preparation

The three datasets are organized around user identifiers to establish relationships between account information, transaction histories, and saved beneficiaries. The user identifier links each transaction to its corresponding primary account holder, while the saved-contact records associate primary users with contact users represented in the users dataset. Transaction identifiers distinguish individual transaction records, and contact identifiers distinguish individual saved-contact relationships.

IV. PROPOSED SYSTEM

The "AI-Powered Multilingual Voice Banking Assistant" is a web-based banking voice app allowing users to conduct chosen banking transactions through voice commands. The design of the application includes such features as speech recognition via a browser, multilingual support, keyword intent recognition, back-end processing, database management, and Text-to-Speech (TTS). The implemented prototype supports English, Hindi, and Marathi languages.

The entire process of communication is done via a workflow approach. First, the user gives the command verbally through the microphone. The Web Speech API of the browser then uses the SpeechRecognition interface to convert this verbal input into text form. The language used for recognition can be set based on the user selection. These languages include English (en-US), Hindi (hi-IN), or Marathi (mr-IN). The processed text is then checked by the application's intent-processing mechanism to determine the

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

banking task that the user intends to do. The corresponding request is then sent to the backend, where it performs the task of communicating with the database.

A. Speech Recognition

Speech recognition constitutes the first step in the implementation of the system. It is performed via the use of the Web Speech API using the browser's SpeechRecognition interface. In this implementation, the Web Speech API is used as either window.SpeechRecognition or window.webkitSpeechReco-
gnition, depending on the supported browser. The language of recognition depends on the language chosen by the user as the interface language. For English, the recognition language is en-US, while for Hindi and Marathi, the languages are hi-IN and mr-IN, respectively.

Voice user interface is an extension rather than replacement of the graphical user interface. User can opt to either give voice commands or operate the controls on the screen. This is because the banking services have already been integrated into the prototype application software.

However, it should be noted that the system only carries out speech recognition and does not do speech identification or biometric verification of the person speaking. This means that voice recognition is used only as an interface tool and not for identifying/authenticating the user.

B. Multilingual Processing

The developed system uses three languages – English, Hindi, and Marathi. Language choice determines speech recognition settings and appropriate response from the user interface.

The system enables users to make equivalent banking inquiries in any of these languages. For instance, a user may issue a request for balance inquiry in English, Hindi, or Marathi. Subsequently, the detected text will be analyzed to detect equivalent banking action.

It allows using one mechanism for voice interactions regardless of user language choice without altering the banking process itself.

Moreover, choosing the language ensures that there is consistency between the input made by the user and the response by the system. In terms of consistency of interaction process in the three languages, the program ties up each language with its own recognition properties and response format. However, variations can be found due to pronunciation, noise, and worded command.

C. Natural Language Processing

The system utilizes a predefined keyword-based intent identification approach to identify the banking operation performed by the user. Once the speech has been translated into text, the application parses the parsed text and matches it against predefined keywords and phrases related to supported banking operations.

Implemented intents include checking the balance, checking transactions, checking account details, performing transfer operations, making payments to beneficiaries, and other banking navigation services. Various phrases related to the same banking operation are mapped to the particular intent.

For instance, phrases like "check my balance," "□□□□ □□□□□□ □□□□," and "□□□□ □□□□□□ □□□□□" are related to the checking balance feature.

The recognized intent triggers a particular predefined operation within the application. Such an approach makes the processing system simple in relation to the banking operations that can currently be performed. The problem is that some commands cannot be properly mapped because they do not correspond to any predefined words or phrases.

Hence, the NLP engine in the current prototype can be regarded as a rule-based and keyword-based intent identification system, as opposed to the intent classification system using machine learning algorithms.

D. Voice Access and Interaction

There exists an interaction technique whereby the user is able to make a supported banking transaction through speech commands. This interaction is integrated with the graphical banking interface and the user is thus able to give speech commands as well as use the banking pages.

The command that is recognized is then translated into a banking function by the application.

E. Backend Processing and Database

Once the intent of the user has been detected, the appropriate request is then processed at the backend of the application. The backend interfaces with the database of the bank to obtain or modify the information that will be needed to process the request.

The database stores information on users, contacts, and transactions that are necessary for the prototype to operate. Balance check requests get account information from the database, while transaction history requests get their transaction records. The transfer and beneficiary operations utilize the information stored on the user and transaction records to complete the request.

F. Text to Speech

Text to Speech will provide the verbal response to the user after processing the requested operation. This verbal response will be provided to the user via the voice interface functionality within the application.

With speech recognition and speech output capabilities, users will be able to interact using voice rather than text only.

G. System Implementation

In the implemented solution, the entire process is done from browser-based speech recognition to backend processing and then the generation of voice response in the process of banking.

The entire process can be depicted as follows:

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

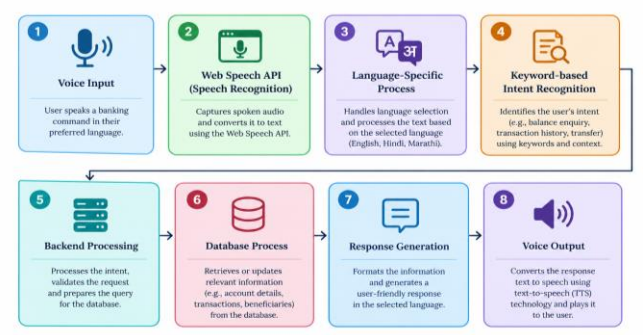


Fig. 01: System Workflow

The system supports basic banking activities that include balance inquiry, transaction history, account detail, money transfer, and payment to beneficiaries.

V. SYSTEM ARCHITECTURE

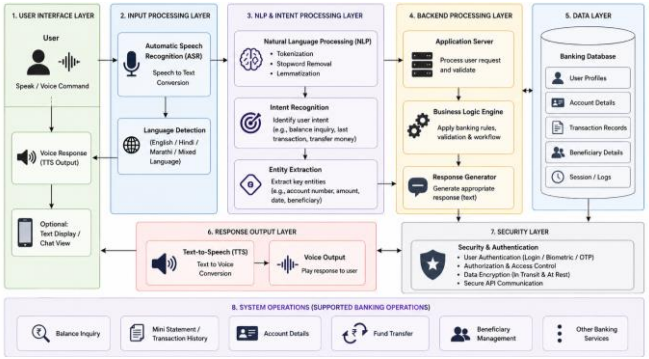


Fig. 02: System Architecture

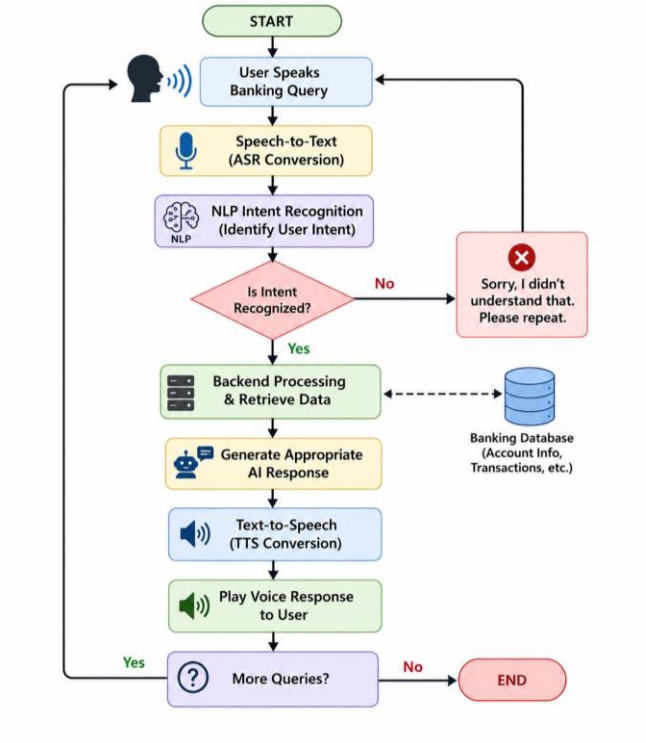


Fig. 03: Flowchart

VI. COMPARATIVE ANALYSIS

In this section, we will compare the proposed model of AI-Powered Multilingual Voice Banking Assistant with other banking models such as branch-based banking and conventional mobile banking apps. This comparison will be made based on the mode of interaction used as well as the features that have been implemented by the system. Traditional banking is typically characterized by human interactions, while conventional mobile banking apps use graphical user interface interactions.

Even if the two models of banking systems have similar objectives, they differ in how users interact with the services. In case of traditional banking systems, bank personnel serves as the medium of interaction. However, graphical user interface is used in mobile banking systems. Our banking system model takes advantage of both graphical user interface and voice recognition interface to access banking operations via voice commands. Their interactions vary in language, navigation, features, and other aspects as indicated in Table 03.

Table 03:Comparative Analysis Table

Feature	Traditional Banking	Mobile Banking Applications	Proposed System
Primary Interaction	Physical interaction with bank staff	Touch and graphical interface	Voice commands with graphical interface
Input Method	Verbal or written interaction with staff	Typing, tapping, and navigation	Spoken commands
Language Support	Depends on staff/service availability	Primarily interface-dependent	English, Hindi, and Marathi
Manual Navigation	Staff-assisted	Required	Reduced for supported voice operations
Voice Interaction	Not applicable	Generally limited	Integrated
Basic Banking Operations	Available through staff	Available through application	Available through voice-assisted workflow
Response Mechanism	Bank staff	Visual interface	Voice and visual response
Security Mechanisms	Institutional banking security	Application and banking security	Prototype-level authentication

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

			; additional transaction security required

VII. IMPLEMENTATION

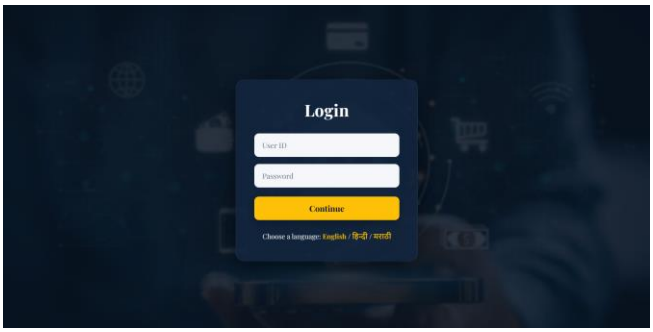


Fig. 04: Login Page

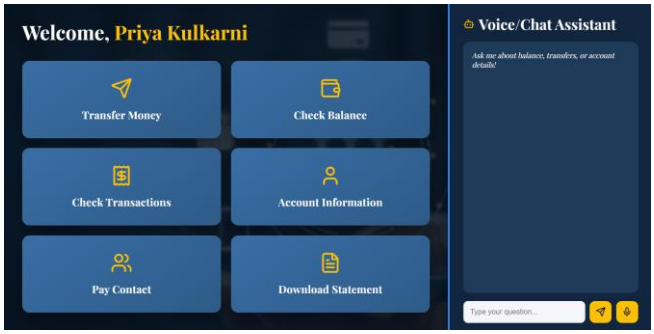


Fig. 05: Dashboard Page (English)



Fig. 06: Dashboard Page (Hindi)



Fig. 07: Dashboard Page (Marathi)

VIII. RESULT AND DISCUSSION

Evaluation of the designed AI-Powered Multilingual Voice Banking Assistant was conducted via prototype-based functional testing to evaluate the performance of voice interactions and bank transactions. Because the designed system is an integrated system and not a machine learning model trained based on a different dataset, the evaluation was mainly concerned with functional implementation of banking activities utilizing the user accounts and transactions provided in the application's database.

Banking activities were conducted via the use of the system using various user accounts and by using voice commands in English, Hindi, and Marathi languages.

A. Multilingual Voice Interaction

The prototype testing included accessing user accounts and making use of commands within the bank interface in the three languages used. Equivalent banking commands were done in voice mode in English, Hindi and Marathi languages to see whether the system could recognize the command from the user and initiate the process.

Results of testing showed that most of the voice commands inputted by the users were recognized and executed. Multilingualism allowed the user to do operations in the bank interface through commands other than the English commands. Hindi and Marathi language provisions are another means of communication for users.

It was noted that the efficiency of voice recognition depended on several factors such as pronunciation, accent, clarity and environment. The presence of noise in the environment influenced speech recognition especially when speech was unclear.

B. Intent Identification and Banking Operations

The intent identification model was employed after the speech recognition process, which helped to identify what type of banking task needed to be performed. It is important to state that the intent identification model was tested on different banking intents rather than being used for processing one command per operation, which is common in other approaches.

The algorithm of intent identification was able to correlate almost all requests received by the system with banking actions to be performed. These included such actions as balance check, transaction view, account detail, fund transfer, and beneficiary actions.

After the correct intent identification was performed by the intent-identification model, the backend had to process the request and interact with the application database in order to retrieve the required information.

C. Functional Evaluation

The functionality assessment consisted of assessing the ability of a user to perform a supported banking operation using voice interaction.

The entire workflow consisted of collecting the voice of the user, transforming speech into text, processing the request depending on the chosen language, recognizing the banking intention, performing a backend process, and providing a response via voice interaction.

This shows that the prototype is capable of

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

combining speech recognition, multilingual processing, intention recognition based on NLP, backend processes, database operations, and response generation based on text-to-speech in one banking application.

For the evaluation process, the user accounts and banking records available in the application have been used.

Table 04: Response Time Analysis Across Supported Languages

Language	No. of Tests	Mean Response Time (s)	Minimum (s)	Maximum (s)	Standard Deviation (s)
English	30	5.82	4.4	7.6	0.82
Hindi	30	6.47	4.9	8.3	0.96
Marathi	30	6.91	5.2	9.0	1.01
Overall	**90**	**6.40**	**4.4**	**9.0**	**0.96**

This can be seen from the table shown above for the different times of response in each of the languages that are supported in this prototype. The response time involves all the time taken in the process of voice inputting, speech recognition, intent detection, backend processing, database transactions, and the formation of the response by the system. From the above results, there is a relatively high response time delay in the prototype, especially during responses in Marathi language compared to English and Hindi language. There could be various factors for this delay, including speech recognition, network services, and voice inputting.

These records have been used as application data for functional testing and not as machine learning datasets for training and testing purposes.

D. User Interaction and Usability

The findings of the prototype evaluation indicate that the voice-first interaction approach could allow reducing the necessity for typing and navigation in order to perform the supported actions. The voice-first approach helped the users to initiate the supported actions via voice and receive voice answers.

The support of English, Hindi, and Marathi provides flexible interaction with the system and is also an alternative way to interact with the text-based bank interface for the speakers of these languages. This is just a prototype evaluation and cannot be viewed as a statistically validated usability or accessibility test.

A usability test specifically including participants from rural areas and elderly people has not been conducted during the present evaluation process. It means that all possible advantages of the system for the mentioned user groups should be regarded as a design goal rather than an experimental finding.

E. Limitations Observed During Evaluation

There are a number of issues raised during the evaluation of this concept. For example, one of the weaknesses is that the performance of the speech recognizer was affected by background noise and differences in speech articulation and accent.

On top of that, although the current prototype is suitable for performing banking operations in a pre-programmed manner, it might be problematic in handling ambiguous and unrestricted requests.

However, apart from the mentioned problems, it should also be noted that although the existing prototype proves the viability of the voice-based banking operations, there are still certain steps that should be taken in order to guarantee the security of this operation in practice. These steps may include authentication of users and transactions, as well as prevention of the use of unauthorized voice commands and banking data.

Therefore, the prototype-based evaluation demonstrates the viability of the multilingual voice-based interface for performing basic banking operations. The successful performance of the tested operations in English, Hindi, and Marathi proves that the created prototype can become the foundation for the development of the multilingual banking interfaces in the future.

F. Security Considerations

In view of the prototype being implemented which includes banking functionalities like money transfer and payment of beneficiaries, security becomes an important issue in the application design. For the time being, security is ensured using the authentication scheme used by the application for user access. Voice input is used as the interface but not for the purposes of authentication.

Currently, there are no features for biometric voice verification, advanced transaction authorization, or any other security feature to ensure detection of fraud in case of unauthorized voice-based commands or inputs. Thus, at this point in time, the implementation should be treated as a research prototype and not as a production banking application.

In order to make the application suitable for use in a real-life banking environment, certain security features have to be added to ensure the safety of the banking transactions and prevent leakage of sensitive banking information.

IX. CONCLUSION

This AI-Powered Multilingual Voice Banking Assistant provides a proof of concept for a multilingual voice interface to be used for performing basic banking operations. The developed prototype is capable of performing speech recognition, multilingual processing, keyword-based intent recognition, backend processing, database processing, and text-to-speech to provide a voice interface to the supported banking operations.

The prototype supports English, Hindi, and Marathi languages and allows users to perform banking operations

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

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

like balance checking, transaction history, account details, money transfer, and pay beneficiaries using voice commands. Functional testing of the prototype on the supported languages revealed that most of the supported voice commands and banking operations were executed successfully.

The multilingual voice interface is created to be used as an alternative way to communicate with the bank through the interface that is different from the text one, especially since the system will be useful for people experiencing problems with typing, reading, or even using graphical interfaces. However, in the current testing phase, there was no special usability test for rural and elderly users.

As a prototype for research purposes, the current version of the system needs some improvements in terms of its security. Further work may include adding support for more Indian languages, implementation of more banking operations, authentication and authorization procedures, noise filtering and robustness, advanced conversational abilities, and full security verification.

Thus, the current system prototype presents the potential of multilingual voice technology as an alternative interaction method.

REFERENCES

[1] R. Sawant, "Analysis of Rural Customers' Dilemmas with Digital Banking Services in Indian Banking Systems: Study of Selected Villages in Pune District of Maharashtra State," *Academy of Marketing Studies Journal*, vol. 27, no. S1, pp. 1–17, 2022.

[2] Y. Mei, "Exploring the mechanisms driving elderly Fintech engagement: the role of social influence and the elderly's digital literacy," *Frontiers in Psychology*, vol. 15, 2024, Art. no. 1420147, doi: 10.3389/fpsyg.2024.1420147.

[3] A. Yadav and J. Patel, "Interactive Voice Response Based Financial Services Solution for Mobile Banking in Rural India," *International Journal of Electrical, Electronics and Computer Systems*, vol. 4, no. 1, 2016, pp. 52–58.

[4] A. Diwan, R. Vaideeswaran, S. Shah, A. Singh, S. Raghavan, S. Khare, V. Unni, S. Vyas, A. Rajpuria, C. Yarra, A. Mittal, P. K. Ghosh, P. Jyothi, K. Bali, V. Seshadri, S. Sitaram, S. Bharadwaj, J. Nanavati, R. Nanavati, and K. Sankaranarayanan, "MUCS 2021: Multilingual and Code-Switching ASR Challenges for Low Resource Indian Languages," in *Proc. Interspeech 2021*, Brno, Czech Republic, 2021, pp. 2446–2450, doi: 10.21437/Interspeech.2021-1339.

[5] H. Palivela, M. Narvekar, D. Asirvatham, S. Bhushan, V. Rishiwal, and U. Agarwal, "Code-Switching ASR for Low-Resource Indic Languages: A Hindi-Marathi Case Study," *IEEE Access*, vol. 13, pp. 9171–9198, 2025, doi: 10.1109/ACCESS.2025.3527745.

[6] N. R. Nithya, S. Malavika, F. Jordan, A. Gangwar, M. N. J. Metilda, S. Umesh, R. Sarab, A. K. Dubey, G. Divakaran, S. V. K. Samudra, and S. V. Gangashetty, "SPRING-INX: A Multilingual Indian Language Speech Corpus by SPRING Lab, IIT Madras," 2023, arXiv:2310.14654.

[7] M. S. Bhatia and S. Khetarpaul, "AI-enhanced bilingual banking assistant," *Scientific Reports*, vol. 15, Art. no. 37526, 2025, doi: 10.1038/s41598-025-22569-z.

[8] C. Bălan, "Chatbots and Voice Assistants: Digital Transformers of the Company–Customer Interface—A Systematic Review of the Business Research Literature," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 18, no. 2, pp. 995–1019, 2023, doi: 10.3390/jtaer18020051.

[9] S. Saranya, J. Jency, V. Aparna, P. Dharshini, and S. Lavanya, "AI Powered Voice Assistant for Banking System," *International Journal of Scientific Research in Science and Technology*, vol. 12, no. 2, pp. 111–119, 2025, doi: 10.32628/IJSRST25122203.

[10] D. Jones, K. Hapeshi, and C. Frankish, "Design guidelines for speech recognition interfaces," Applied Ergonomics, vol. 20, no. 1, pp. 47–52, 1989, doi: 10.1016/0003-6870(89)90009-4

[11] M. Wolters, K. Georgila, J. D. Moore, and S. E. MacPherson, "Being Old Doesn't Mean Acting Old: How Older Users Interact with Spoken Dialog Systems," ACM Transactions on Accessible Computing, vol. 2, no. 1, Article 2, pp. 1–39, 2009, doi: 10.1145/1525840.1525842

[12] B. Stigall and K. Caine, "A Systematic Review of Human Factors Literature About Voice User Interfaces and Older Adults," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 64, no. 1, 2020, published online Feb. 9, 2021, doi: 10.1177/1071181320641004

[13] R. Eustaquio-Jiménez, M. Durand-Azurza, J. Gamboa-Cruzado, M. León Morales, N. Pajares Ruiz, R. López de Montoya, and N. M. Ortiz Céliz, "Chatbots for customer service in financial entities—A comprehensive systematic review," Journal of Infrastructure Policy and Development, vol. 8, no. 16, Art. no. 10122, 2024, doi: 10.24294/jipd10122

[14] G. Graham, T. M. Nisar, G. Prabhakar, R. Meriton, and S. Malik, "Chatbots in customer service within banking and finance: Do chatbots herald the start of an AI revolution in the corporate world?" Computers in Human Behavior, vol. 165, Art. no. 108570, 2025, doi: 10.1016/j.chb.2025.108570

[15] M. M. Mariani, N. Hashemi, and J. Wirtz, "Artificial intelligence empowered conversational agents: A systematic literature review and research agenda," Journal of Business Research, vol. 161, Art. no. 113838, 2023, doi: 10.1016/j.jbusres.2023.113838

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