IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 11, November 2026

AUTOMATED BANKING TRANSACTION SYSTEM VIA VOICE RECOGNITION

Kartik Ravindra Kakuste[1], Prof. Manoj V. Nikum[2], Prof. Pravin I. Patil[3]

Research Scholar MCA, Shri Jaykumar Rawal Institute of Technology, Dondaicha, India[1]

HOD MCA & Assistant Professor, Shri Jaykumar Rawal Institute of Technology, Dondaicha, India[2]

Assistant Professor, Shri Jaykumar Rawal Institute of Technology, Dondaicha, India[3]

Abstract: This paper presents the proposed design of an Automated Banking Transaction System via Voice Recognition, an academic prototype intended to support selected banking operations through spoken commands. Automatic Speech Recognition (ASR) converts the user's speech into text, while Natural Language Processing (NLP) identifies the requested operation and extracts relevant details. The proposed functions include balance inquiry, mini statement retrieval and a controlled fund-transfer workflow. User verification, input validation and explicit confirmation are included as safeguards before sensitive actions. The prototype is intended to use fictional account data and simulated transactions rather than connect to a live banking network. The evaluation plan considers recognition accuracy, intent interpretation, parameter extraction, response time, task completion and safe rejection of invalid requests. Measured results will be added after implementation and testing.

Keywords: Voice Recognition, Automated Banking, Automatic Speech Recognition, Natural Language Processing, Voice Authentication, Fund Transfer.

I. I. INTRODUCTION

Digital banking provides customers with access to account information and transaction services through electronic interfaces. Most conventional interfaces depend on visual navigation, menus and manual text entry. A voice-enabled interface provides an alternative interaction method in which users can express a banking request verbally.

A voice interface must do more than recognize words. It must identify the user's intention, extract details such as a transaction amount or beneficiary, verify the user and ensure that an operation is valid before it is executed. A speech-recognition error may change an amount or select the wrong operation; therefore, validation and confirmation are essential.

This paper proposes a voice-enabled banking interface combining ASR, NLP and a controlled banking-service layer. Its intended functions are balance inquiry, mini statement retrieval and simulated fund transfer. The academic prototype uses fictional accounts and transactions and is not connected to a live banking network.

II. LITERATURE REVIEW

Voice-enabled banking relies on speech processing, language understanding and security mechanisms. ASR converts audio input into text, while NLP identifies the intended operation and extracts relevant entities. Different phrases may express the same intent; for example, “check my balance” and “what is my current balance?” can both represent a balance inquiry.

Voice biometrics may support speaker verification, but voice matching alone should not be treated as sufficient protection for high-risk operations. Multi-factor authentication and explicit transaction confirmation can provide additional safeguards. The supplied project references discuss ASR, NLP, voice biometrics and deep-learning approaches; these materials provide background concepts rather than measured results from this implementation.

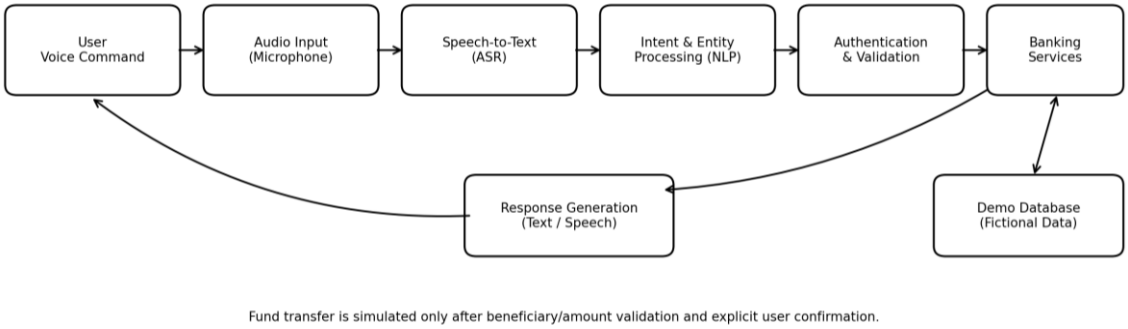
A. A. Research Gap and Project Scope

The project focuses on integrating spoken commands with a limited set of banking workflows and documenting the validation process. The system is intended to demonstrate balance inquiry, mini statement retrieval and simulated fund transfer using synthetic data. It does not claim to replace commercial banking systems or provide real payment processing.

III. METHODOLOGY

The proposed system is organized into audio capture, speech-to-text conversion, intent interpretation, user verification, banking-service processing, database access and response generation. The modules separate speech understanding from transaction execution.

Fig. 1. Proposed System Architecture



As shown in Fig. 1, the spoken request passes through audio capture, ASR, NLP, authentication and validation before the banking-service layer accesses the fictional demonstration database. The response is returned as text or speech. Sensitive transfer requests remain subject to explicit confirmation.

B. A. Voice Input and Intent Processing

The audio-capture module receives the user's speech. The ASR module converts the audio signal into text. The NLP module classifies the request into a supported intent and extracts entities such as an amount, beneficiary or requested transaction count. If the command is unclear, incomplete or unsupported, the system asks the user to clarify it instead of continuing.

C. B. Banking Operations

Balance inquiry retrieves the available balance for the authenticated demonstration account. Mini statement retrieves a limited number of recent synthetic transactions. Fund transfer collects the beneficiary and amount, validates the details, summarizes the request and asks for explicit confirmation. Only a confirmed request is recorded as a simulated transaction.

D. C. Authentication and Security

The user must pass the configured verification process before account-specific information is accessed. For sensitive operations, identity verification is followed by a separate confirmation step. Voice matching alone should not be treated as sufficient protection. Invalid amounts, unknown beneficiaries, failed authentication and cancelled requests must be rejected without changing account records.

E. D. System Data and Processing

The demonstration database may contain user profiles, account records, beneficiary details and transaction history. All records must be fictional test data. Sensitive values should be masked where possible, and credentials or one-time passwords must not be stored in plain text. The prototype must not connect to live banking services or process real funds.

IV. SYSTEM WORKFLOW AND VALIDATION DETAILS

F. A. Command Processing and Transaction States

The command-processing stage maps recognized text to supported intents such as BALANCE_INQUIRY, MINI_STATEMENT, FUND_TRANSFER, CANCEL_TRANSFER and REPEAT_RESPONSE. Entity extraction identifies the values required by each intent. If an amount or beneficiary is missing, the application asks a follow-up question. An uncertain intent is not passed to the banking-service layer.

A transfer request should move through explicit states: initiated, details-collected, validated, awaiting-confirmation, confirmed-simulated, cancelled or rejected. The state model prevents a recognized sentence from directly updating the account balance. A transaction is recorded only after the user reviews the summary and confirms it. Cancellation before confirmation must leave the demonstration account unchanged.

G. B. Proposed Test Scenarios

Positive scenarios include a clear balance request, an alternate balance phrase, a mini-statement request and a transfer containing a valid beneficiary and amount. Negative scenarios include empty audio, an unsupported request, missing transfer details, an unknown beneficiary, a non-positive amount, authentication failure and cancellation before confirmation.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 11, November 2026

TABLE I. PROPOSED FUNCTIONAL TEST SCENARIOS

Test ID	Input scenario	Expected outcome
TC-01	Clear balance inquiry	Return demo balance after authentication
TC-02	Mini statement request	Return recent synthetic transactions
TC-03	Transfer with valid details	Validate and request confirmation
TC-04	Missing amount / beneficiary	Ask for missing information
TC-05	Unknown beneficiary / invalid amount	Reject without database update
TC-06	User cancels pending transfer	Cancel; no transaction recorded
TC-07	Unsupported or unclear speech	Ask user to repeat or clarify
TC-08	Authentication failure	Do not expose account information

H. C. Evaluation Metrics

Recognition accuracy measures the proportion of voice commands transcribed correctly. Intent accuracy measures whether the system identifies the expected banking operation. Entity-extraction accuracy measures whether required values, such as amount and beneficiary, are extracted correctly. Average response time measures the elapsed time from the end of the spoken command to the system response. Task-completion rate measures the proportion of valid requests that reach the correct outcome.

The evaluation should record the recognized text, expected intent, predicted intent, extracted parameters, response time and final outcome for each test. Results should be reported separately for each operation and test condition.

V. RESULT AND DISCUSSION

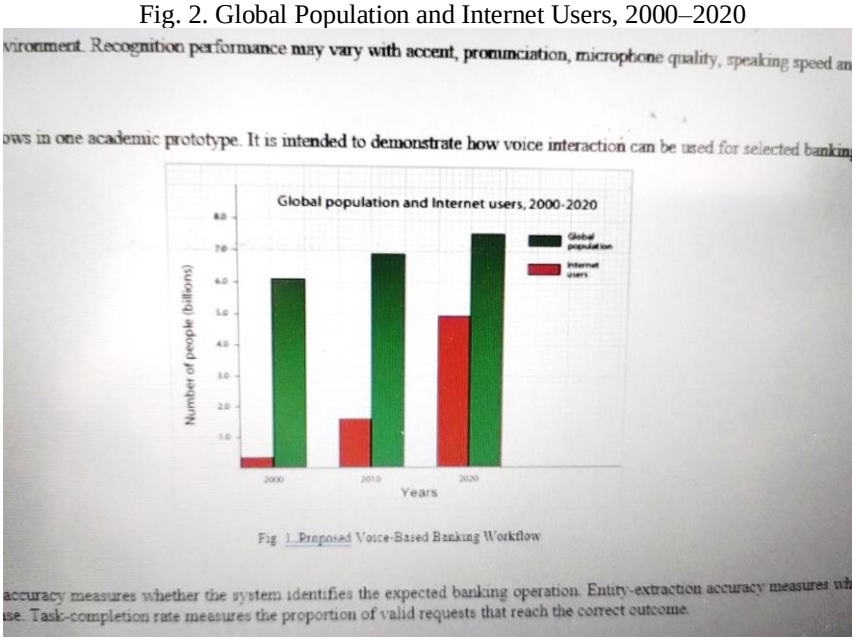
The implementation and experimental validation are pending. Therefore, this paper does not claim a recognition-accuracy percentage, intent-classification score or response-time value. The planned evaluation will use a labelled command set covering balance inquiry, mini statement, fund transfer, cancellation, unclear speech and unsupported requests.

The principal measures are speech-recognition accuracy, intent accuracy, entity-extraction accuracy, average response time, task-completion rate and safe-rejection rate. Actual values must be populated only after tests are executed. Figures from background project materials must not be represented as results obtained by this project.

I. A. Limitations and Security Considerations

Recognition quality may vary with accent, pronunciation, speaking speed, microphone characteristics and background noise. Voice-based verification may also be affected by replay or impersonation attempts and should not be the sole control for sensitive operations. The prototype is limited to a small set of supported commands and simulated banking data.

Fig. 2. Global Population and Internet Users, 2000–2020



The figure is included as contextual background in the limitations and scope discussion. It illustrates the broad growth of internet usage alongside global population over the period shown in the supplied chart. It is not a measurement or result generated by the proposed banking prototype.

The system is intended as an academic prototype, not a production banking channel. It must not store real banking credentials, connect to live accounts or transfer real funds. Any future integration with banking infrastructure would require formal authorization, security assessment and compliance with applicable requirements.

VI. CONCLUSION

This paper presents the proposed design of an Automated Banking Transaction System via Voice Recognition. The system combines ASR, NLP, user verification and a controlled banking-service layer to support balance inquiry, mini statement retrieval and simulated fund transfer. The design treats speech recognition as the first step in a validated workflow rather than as direct authorization to perform a transaction.

Future work may include improving recognition across accents and languages, adding more banking intents, evaluating speaker verification with additional authentication factors, improving privacy controls for spoken responses and conducting structured usability testing. Actual performance findings will be added only after implementation and testing.

ACKNOWLEDGMENT

The author acknowledges the academic guidance of **Prof. Pravin I. Patil** and the support and facilities provided by the Department of MCA, Shri Jaykumar Rawal Institute of Technology, Dondaicha.

REFERENCES

[1] “Automated Banking Transaction System via Voice Recognition,” project research paper supplied as reference material.
[2] “Automated Banking System via Voice Recognition – Revolutionizing Financial Accessibility through AI,” project presentation supplied as reference material.
[3] “Automated Banking Transaction System via Voice Recognition,” project abstract supplied as reference material.
[4] A. Graves, A.-r. Mohamed and G. Hinton, “Speech Recognition with Deep Recurrent Neural Networks,” Proc. IEEE ICASSP, pp. 6645–6649, 2013.
[5] Python Software Foundation, “Python Documentation,” https://docs.python.org/.
[6] TensorFlow, “TensorFlow Documentation,” https://www.tensorflow.org/.