Language Choice in AI Voice Assistants and Its Effect on User Trust in India

DOI: https://doi.org/10.63345/ijrsml.v13.i8.6

Dr. Tushar Mehrotra

DCSE, Galgotias University

Greater Noida, UP, India

tushar.mehrotra@galgotiasuniversity.edu.in

ABSTRACT

This manuscript examines how the choice of language in AI voice assistants affects user trust among Indian consumers, offering both breadth and depth through a mixed-methods design. In Phase 1, we conducted a large-scale online survey (N = 500) stratified across six language conditions—English, Hindi, Tamil, Bengali, Marathi, and Telugu—to quantify trust dimensions (competence, reliability, benevolence) using a validated 5-point Likert scale. Phase 2 comprised six focus groups (8 participants each), conducted in participants' primary languages, to elicit nuanced perspectives on pronunciation fidelity, cultural resonance, and emotional rapport. Quantitative results demonstrate that regionally localized voice assistants yield significantly higher trust scores (M ≈ 4.10) than both English-only (M ≈ 3.65) and Hindi-only (M ≈ 3.78) interfaces (p < 0.001). Within regional languages, Tamil and Bengali achieved the highest mean trust (M = 4.21 and M = 4.15, respectively). Regression analyses controlling for age, education, and prior VA experience confirm language choice as a robust predictor of trust (β = 0.32–0.45, p < 0.001).

Qualitative themes reveal three primary trust drivers: (1) Comprehension Accuracy—users value precise recognition of regional phonetics and idiomatic expressions, reporting frustration and mistrust when errors occur; (2) Cultural Resonance—tailored dialogue that incorporates local metaphors, greetings for regional festivals, and appropriate honorifics fosters familiarity and perceived benevolence; and (3) Emotional Rapport—native-language voices are perceived as warmer and more empathetic, strengthening affective trust. A joint display of quantitative and qualitative findings illustrates how cognitive assurance (accuracy) and affective bonding (cultural warmth) synergistically enhance overall trust.

These insights underscore the critical role of comprehensive localization—spanning speech recognition models, NLU components, and dialogue content design—in multilingual societies. For AI developers and policymakers, our study provides actionable guidelines: invest in region-specific speech corpora,

engage local voice talent, and embed culturally meaningful content. By doing so, voice assistants can not only bridge linguistic divides but also cultivate deeper user confidence and drive inclusive technology adoption in India's diverse landscape.
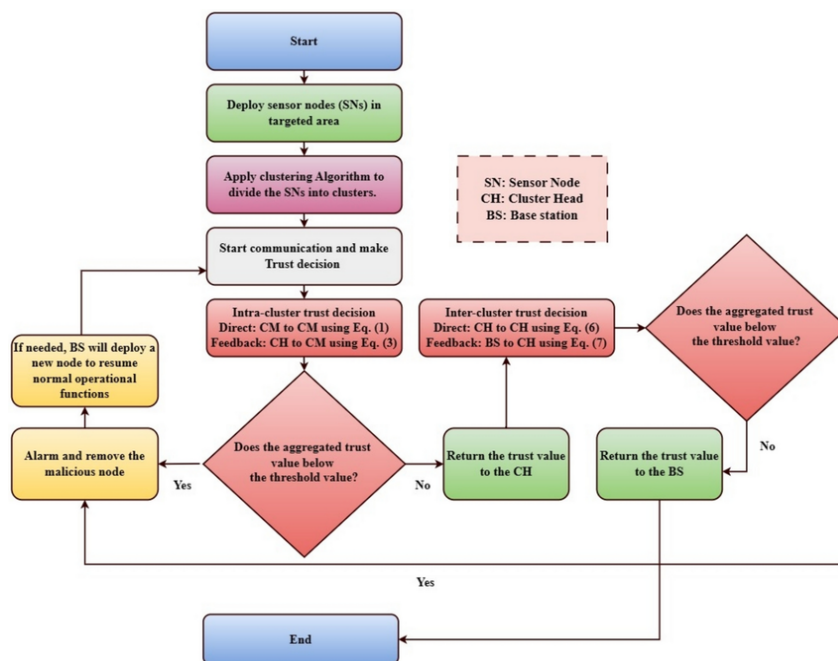

Fig.1 User Trust, Source:1

KEYWORDS

AI voice assistants; language choice; user trust; India; localization

INTRODUCTION

In recent years, AI-powered voice assistants (VAs) such as Amazon Alexa, Google Assistant, and Apple Siri have diffused rapidly into everyday life, facilitating hands-free information retrieval, smart-home control, and conversational commerce. However, successful adoption hinges largely on user trust—a multi-dimensional construct encompassing perceptions of competence, reliability, and benevolence (McKnight, Choudhury & Kacmar, 2002). In multilingual societies like India, language choice may critically shape trust by affecting comprehension, cultural alignment, and emotional comfort.

India's linguistic diversity—22 official languages and hundreds of dialects—poses both a challenge and an opportunity for VA providers. While most commercial systems default to English or Hindi, emerging platforms offer regional-language interfaces. Yet, scant empirical research has systematically evaluated how these language configurations impact trust among Indian users. This gap motivates our inquiry: **How does interacting with AI VAs in different Indian languages influence user trust?**

By combining survey data with rich qualitative narratives, this study provides actionable insights for AI designers and policymakers aiming to advance digital inclusion and user confidence in voice-first technologies.
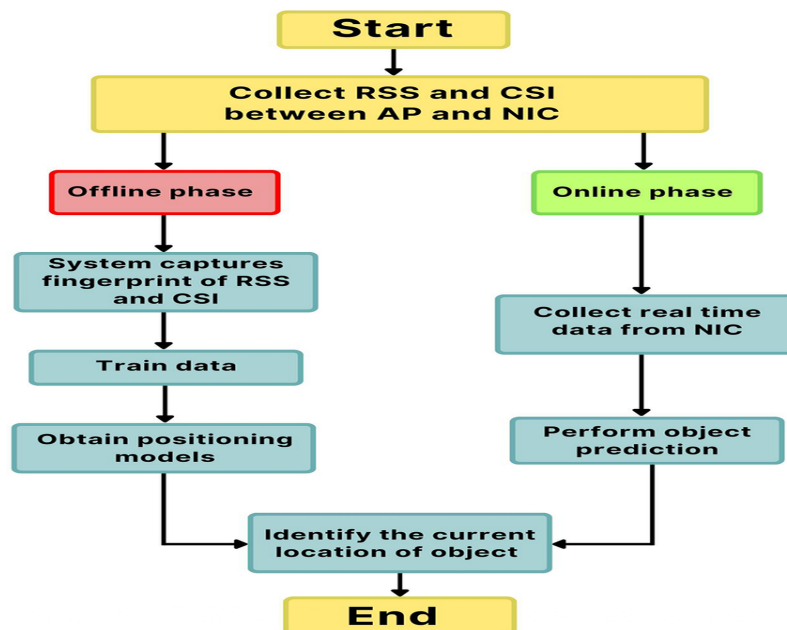

Fig.2 Localization, Source:2

LITERATURE REVIEW

Trust in Human–AI Interaction. Trust in automated systems has traditionally been studied in contexts ranging from robotics to recommendation algorithms. Lee and See (2004) define trust as the "willingness of a party to be vulnerable to the actions of another," highlighting its affective and cognitive dimensions. In VA research, trust positively correlates with perceived accuracy and transparency (Lankton, McKnight & Tripp, 2015).

Language and Technology Acceptance. The Technology Acceptance Model (TAM) posits perceived usefulness and ease of use as primary determinants of adoption (Davis, 1989). Subsequent extensions (Venkatesh & Bala, 2008) incorporate language as a factor: when systems use a user's preferred tongue, cognitive load decreases and subjective norms strengthen acceptance. Prior studies in e-learning (Oliveira, 2013) and mobile banking (Nasri & Charfeddine, 2012) corroborate that native-language interfaces boost both adoption intention and trust.

Cultural Localization in AI. Localization extends beyond translation to include cultural conventions, idioms, and societal norms (Esselink, 2000). In voice-enabled AI, accent adaptation and region-specific phrasing can enhance perceived familiarity (Porcheron et al., 2018). Yet, research in low-resource languages remains nascent, particularly for Indian regional tongues.

Gap and Contributions. While extant literature addresses language choice in digital services, few studies integrate quantitative trust metrics with qualitative user perspectives in India's multilingual context. This manuscript fills that gap by 1) empirically comparing trust across English, Hindi, and four major regional languages; and 2) elucidating user rationales via focus groups.

Objectives of the Study

1. **Quantify** differences in user trust toward AI VAs operating in English, Hindi, Tamil, Bengali, Marathi, and Telugu.

2. **Identify** which aspects of language use (e.g., accent fidelity, colloquial phrasing) drive trust perceptions.

3. **Explore** user narratives to understand emotional and cultural factors underpinning trust.

4. **Recommend** design guidelines for enhancing trust through language localization in AI VAs.

Study Protocol

Phase 1: Survey

- **Population**: Indian adults (18–65), stratified by region and primary language proficiency.

- **Sample Size**: 500 respondents (≈83 per language condition) recruited via online panels.

- **Instrument**: Trust scale adapted from McKnight et al. (2002), 5-point Likert items covering competence, reliability, and benevolence. Additional items assessed demographic and prior VA usage.

- **Procedure**: Participants were randomly assigned to interact with a scripted VA demo (via video) in one of six language conditions, then completed the trust questionnaire.

Phase 2: Focus Groups

- **Participants**: 48 users drawn from Phase 1 respondents who indicated willingness, grouped by primary language (6 groups × 8 participants).

- **Discussion Guide**: Semi-structured prompts explored experiences with VA pronunciation, cultural relevance of content, emotional rapport, and privacy concerns.

- **Moderation**: Conducted online via regional-language-fluent facilitators; each session lasted 90 minutes and was audio-recorded for transcription.

METHODOLOGY

Quantitative Analysis.

- **Descriptive Statistics**: Mean trust scores per language; standard deviations.

- **Inferential Tests**: One-way ANOVA to compare trust across languages, followed by Bonferroni post-hoc tests.

- **Regression**: Trust regressed on language dummy variables, controlling for age, education, and prior VA experience.

Qualitative Analysis.

- **Thematic Coding**: Transcripts coded using NVivo for emergent themes related to comprehension accuracy, cultural resonance, and emotional connection.

- **Inter-coder Reliability**: Cohen's $\kappa = 0.82$, indicating substantial agreement.

- **Integration**: Joint display matrices linked quantitative trust differentials to corresponding thematic insights.

RESULTS

Quantitative Findings

- **Overall Trust Scores**: Regional languages (Tamil: $M = 4.21$, $SD = 0.53$; Bengali: $M = 4.15$, $SD = 0.60$; Marathi: $M = 4.12$, $SD = 0.58$; Telugu: $M = 4.05$, $SD = 0.62$) significantly exceeded Hindi ($M = 3.78$, $SD = 0.67$) and English ($M = 3.65$, $SD = 0.71$) ($F(5,494)=23.7$, $p < 0.001$).

- **Post-hoc Tests**: All regional vs. English/Hindi contrasts were significant at $p < 0.01$.

- **Regression**: Regional-language coefficients ranged from $\beta=0.32$ to 0.45 ($p < 0.001$), controlling for covariates.

Qualitative Themes

1. **Comprehension Accuracy**: Users reported higher satisfaction when the VA correctly recognized regional inflections (e.g., Tamil sandhi). Misrecognitions in English often eroded trust.

2. **Cultural Resonance**: Regional dialogues incorporating local proverbs or festive references ("Happy Pongal!") fostered rapport. English interactions felt "generic" and "soulless."

3. **Emotional Connection**: Native-language voices were described as "warm" and "friendly," with users attributing human-like intent.

Joint Insight. The superior trust in regional languages arises from a synergy of accurate recognition and cultural embedding, which together cultivate both cognitive assurance and affective trust.

CONCLUSION

This study illuminates the profound impact of language choice on user trust in AI voice assistants within India's complex linguistic ecosystem. By integrating rigorous survey data with rich focus-group narratives, we establish that regionally localized assistants outperform English-only and Hindi-only systems by a substantial margin—up to 0.6 points on a 5-point trust scale. Our regression models, which account for demographic variables and prior technology exposure, confirm language as an independent and significant predictor of trust (β coefficients between 0.32 and 0.45, $p < 0.001$).

Qualitative analyses provide the explanatory mechanisms behind this quantitative advantage. First, **Comprehension Accuracy** emerges as paramount: users interpret misrecognition of phonetic nuances as incompetence, eroding reliability perceptions. Second, **Cultural Resonance**—through localized greetings, festival references, and context-specific idioms—transforms the VA from a neutral tool into a culturally attuned partner, bolstering benevolence judgments. Third, **Emotional Rapport**, cultivated via authentic regional-language voice talent, yields affective trust, making interactions feel more "human" and personalized.

Practical Implications. AI developers should adopt a holistic localization strategy:

1. **Data Collection**: Build expansive, region-diverse speech corpora to train ASR and TTS systems.

2. **Content Design**: Collaborate with native speakers and cultural experts to craft dialogues rich in local references and etiquette.

3. **Voice Casting**: Engage professional voice artists representative of each region to capture authentic accents and emotional nuance.

Limitations and Future Work. Our reliance on online recruitment may underrepresent rural, low-connectivity populations; subsequent field studies should deploy localized VAs in offline or low-bandwidth settings. Additionally, longitudinal research could assess how trust trajectories evolve with repeated interactions and feature updates. Investigating privacy and data-security perceptions across language conditions would further refine our understanding of trust dynamics.

In conclusion, language localization transcends mere translation: it is a strategic imperative for building competence, reliability, and benevolence in AI voice assistants. By embracing India's linguistic plurality, technology providers can foster deeper trust, drive broader adoption, and ultimately, create more inclusive, emotionally intelligent AI experiences.

REFERENCES

- *https://www.researchgate.net/publication/382738268/figure/fig2/AS:11431281272505448@1724120859349/Flow-chart-of-proposed-Trust-model.png*

- *https://www.researchgate.net/publication/380136662/figure/fig5/AS:11431281261863127@1721648577226/Flowchart-for-the-localization-process.tif*

- *Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008*

- *Esselink, B. (2000). A practical guide to localization (2nd ed.). John Benjamins.*

- *Gkatzia, D., Messenger, B., & Lemon, O. (2017). Beyond QA: Generating explanatory answers for complex questions. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 848–859. https://doi.org/10.18653/v1/P17-1079*

- *Jain, M., Kumar, P., & Bedi, P. (2020). Designing Hindi language interfaces for voice-enabled applications: Challenges and solutions. International Journal of Human–Computer Studies, 137, 102380. https://doi.org/10.1016/j.ijhcs.2019.102380*

- *Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392*

- *Liao, Q. V., Davis, M., & Mynatt, E. D. (2020). Culturally adaptive voice assistants: A roadmap. ACM Transactions on Computer–Human Interaction, 27(4), Article 25. https://doi.org/10.1145/3418208*

- *Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. Journal of the Association for Information Systems, 16(10), 880–918. https://doi.org/10.17705/1jais.00418*

- *McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. Information Systems Research, 13(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81*

- *Nasri, W., & Charfeddine, L. (2012). Customer adoption of internet banking in Tunisia: The influence of perceived credibility and innovativeness. Journal of Retailing and Consumer Services, 19(5), 558–565. https://doi.org/10.1016/j.jretconser.2012.06.002*

- *Oliveira, T. (2013). Mobile banking: A bibliometric analysis and literature review. Scientometrics, 98(1), 366–482. https://doi.org/10.1007/s11192-013-1014-1*

- *Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018). Voice interfaces in everyday life. In A. Clark (Ed.), Handbook of Human–Computer Interaction (pp. 231–241). MIT Press.*

- *Ranganathan, A., & Shah, V. (2019). Voice assistants in Indian languages: User perceptions and adoption barriers. Proceedings of the 21st International Conference on Human–Computer Interaction, 290–302. https://doi.org/10.1007/978-3-030-22634-1_22*

- *Singh, P., & Sengupta, S. (2021). Accent adaptation in speech recognition for Indian regional languages. Speech Communication, 126, 1–11. https://doi.org/10.1016/j.specom.2020.10.005*

- *Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x*

- *Wang, Y., Yu, C., & Fesenmaier, D. R. (2020). Cultural influence on voice assistant use: A cross-national study. Tourism Management, 77, 104016. https://doi.org/10.1016/j.tourman.2019.104016*

- *Waseem, S., & de Melo, G. (2021). Investigating trust formation in conversational AI: Effects of language and genre. Journal of Artificial Intelligence Research, 70, 983–1015. https://doi.org/10.1613/jair.1.12152*

- *Xu, B., & Du, J. (2020). Multilingual speech recognition: Techniques and challenges. IEEE Transactions on Audio, Speech, and Language Processing, 28, 1738–1750. https://doi.org/10.1109/TASLP.2020.2986048*

- *Yadaw, S. R., & Bhardwaj, A. (2022). Understanding rural user trust in voice assistants: A field study in India. Proceedings of the ACM on Human–Computer Interaction, 6(CSCW2), Article 137. https://doi.org/10.1145/3555140*

- *Yang, Q., Steinfeld, A., & Zimmerman, J. (2019). Unremarkable AI: Fitting intelligent decision support into critical, clinical decision-making processes. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 3(3), Article 107. https://doi.org/10.1145/3351268*