Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/BA08.20250371
Published: December 20, 2025

Benefits of NLP Tools for Speaking Development of Second Language Learners

S. M Mehedi Hasan, Anika Tahmina Chowdhury and Md. Shohaib Islam
Department of Computer Science
American International University-Bangladesh
Dhaka, Bangladesh
22-48983-3@student.aiub.edu, 22-48949-3@student.aiub.edu, 22-49009-3@student.aiub.edu

Iftekhar Mahmud
Department of English
American International University-Bangladesh
Dhaka, Bangladesh
iftekhar@aiub.edu

Abstract

Having to build oral proficiency can be difficult for L2 learners, mainly because they often lacked enough chances to practice and receive feedback. While NLP tools presented potential for automated speaking practice, research on their effectiveness from the learner's viewpoint is still evolving. This study employed a quantitative research framework. Data was collected via Google Forms survey L2 learners using NLP tools (e.g., speech recognition, automated pronunciation evaluation, conversational AI). The survey examined improvements in fluency, precision, self-assurance, and participant contentment. The analysis was expected to reveal a strong positive correlation between NLP tool usage and self-reported speaking skill enhancements. Projected findings showed significant improvements in recognized pronunciation accuracy and spoken fluency. Results also demonstrated that immediate, automated feedback enhanced learner confidence and motivation for oral practice. The outcomes indicated that NLP tools were effective further resources for L2 speaking development. They offered scalable, personalized practice addressing limitations in traditional educational settings. In the long run, applying NLP-based speaking tools into language programs enabled learners to adopt a more proactive and confident approach to achieving oral proficiency.

Keywords
AI-assisted learning, Automated feedback, L2 learning, Oral proficiency, Pronunciation accuracy.

1. Introduction
In today's world, artificial intelligence (AI) is becoming increasingly integrated into education, with digital assistants such as Siri and language learning platforms such as Duolingo playing key roles in everyday life. For second language (L2) learners, speaking skills remain one of the most challenging aspects of language acquisition, often hindered by limited practice opportunities, anxiety, and insufficient, personalized feedback (Litman et al., 2018). Natural Language Processing (NLP), a branch of AI that enables machines to understand and generate human language, has emerged as a promising tool for addressing these challenges. NLP technologies, such as Automatic Speech Recognition (ASR) systems, AI chatbots, and intelligent tutoring systems, are reshaping how speaking skills are developed and practiced. These tools provide real-time feedback, simulate conversational practice, and enhance learning through personalized adaptive learning experiences.

Research on the benefits of NLP tools in L2 education is increasing. For instance, Mobile-Assisted Language Learning (MALL), which often incorporates NLP technologies, has shown significant improvements in speaking skills, particularly in pronunciation accuracy (Li 2024) Similarly, AI chatbots have been found to positively impacted speaking and listening scores while increasing motivation and willingness to communicate (Koç and Savaş, 2024) Tools such as ELSA Speak, which uses ASR for pronunciation training, and ChatGPT, which enables open-ended conversational practice, illustrate how NLP-driven platforms offer accessible, low-anxiety environments for learners to practice speaking. Furthermore, intelligent tutoring systems, such as Duolingo's AI-powered exercises, personalize speaking tasks to meet the individual needs of learners, making practice more engaging and effective (Son et al., 2023).

Despite these advancements, the existing literature reveals the need for a more unified view of the benefits of NLP tools for L2 speaking development. While studies have explored the effectiveness of different tools, such as ASR-based pronunciation trainers (Sun 2023; Bashori et al., 2024) and dialogic AI systems (Bibauw et al., 2022; Du and Daniel, 2024), the findings are often fragmented across subfields like CALL (Computer-Assisted Language Learning (CALL), MALL, and speech technology. This fragmentation makes it difficult to comprehensively understand how NLP technologies collectively support speaking proficiency. Moreover, previous reviews have focused on specific tools or broader methodological critiques without synthesizing the benefits of NLP as a whole (Amrate and Tsai, 2024; Kang et al., 2024).

The significance of this research lies in its potential to inform educators, curriculum developers, and policymakers of the advantages of integrating NLP tools into L2 speaking instruction. Speaking proficiency is a crucial part of communicative competence and is often a source of anxiety and learner attrition (Du and Daniel, 2024). NLP tools can help alleviate these challenges by providing personalized, scalable practice that may not be available in traditional settings. For example, ASR systems enable learners to receive immediate feedback on their pronunciation and fluency (Kang et al., 2024; Sun, 2023) whereas AI chatbots offer opportunities for interactive dialogue in a pressure-free environment (Koç and Savaş, 2024). Additionally, advanced multimodal NLP systems that combine speech, text, and visual cues offer more comprehensive feedback, representing the future of personalized language learning (Dubey et al., 2025).

This study aims to clarify the benefits of NLP technologies for speaking development by synthesizing evidence from various studies, encouraging their informed adoption, and promoting further innovation in this field.

Research Questions

This study addresses the following research questions:

1. How do second-language learners perceive the impact of NLP tools on their speaking fluency and pronunciation accuracy?
2. To what extent do NLP tools increase learner confidence and reduce anxiety when speaking English?
3. How effective are automated feedback systems in helping learners correct their mistakes and improve their speaking skills?

This paper aims to provide a comprehensive overview of the benefits of NLP in speaking instruction by analyzing these questions through a synthesis of recent empirical and review studies on the subject. This emphasizes that NLP is not a singular technology but a versatile framework encompassing various tools, from speech recognition to conversational AI, all designed to enhance the effectiveness, engagement, and accessibility of L2 speaking practice.

2. Literature Review

Consider a second-language learner who feels anxious and hesitant when speaking in class or worries about pronunciation errors; now imagine the same learner using an AI-powered tool that provides immediate, personalized feedback in a stress-free environment. Recent advances in Artificial Intelligence (AI) and Natural Language Processing (NLP) are making such scenarios increasingly possible. This review synthesizes current literature on the benefits of NLP-powered tools for addressing these challenges and is organized into five sections: (1) the speaking condition of L2 learners, (2) general advantages of technology, (3) the role of AI in education, (4) specific applications of NLP, and (5) a focused analysis of AI/NLP tools for speaking development.

2.1 Speaking Condition of Second Language Learners

Second language (L2) learners face psychological, linguistic, and external challenges that hinder speaking proficiency. Anxiety, fear of judgment, and low motivation reduce confidence and willingness to speak (Kashinathan & Aziz, 2021; Mora et al., 2023). Linguistic difficulties include pronunciation errors, limited vocabulary, and mother tongue interference, which affect fluency and pragmatic skills (Xu & Cao, 2022; Bashori et al., 2024). Externally, limited opportunities for meaningful practice and comprehension challenges in synchronous settings further restrict progress (Kashinathan & Aziz, 2021; Qiao & Yijun, 2023). These factors create a cycle where anxiety limits practice, and limited practice perpetuates low proficiency, highlighting the need for safe, effective interventions such as NLP-based tools.

2.2 Advantages of Technology for Speaking Development

Technology, including Mobile-Assisted Language Learning (MALL) and Computer-Assisted Pronunciation Training (CAPT), transforms speaking practice by creating low-anxiety, pressure-free environments (Kashinathan and Aziz, 2021; Sun, 2023). These tools provide immediate, personalized feedback on pronunciation (Amrate and Tsai, 2024; Sun, 2023), enhance engagement, motivation, and self-efficacy (Li, 2024; Dubey et al., 2025), and support flexible, anytime-anywhere learning with features like live transcripts for comprehension (Kashinathan and Aziz, 2021; Qiao and Yijun, 2023).

2.3 AI in Education

The integration of AI in education enables personalized, adaptive, and intelligent learning, particularly enhancing language acquisition. AI systems, such as Intelligent Tutoring Systems and chatbots, tailor instruction and provide real-time feedback to learners' proficiency levels, improving pronunciation and fluency (Kartal and Yeşilyurt, 2024; Son et al., 2023; López-Minotta et al., 2025). AI also diversifies learning tools through gamification, robot-assisted learning, and context-based applications like image-recognition-assisted vocabulary practice (Kartal and Yeşilyurt, 2024; Jia et al., 2022), while automating assessment to offer fast, consistent, and objective feedback. Nonetheless, challenges remain, including algorithmic bias and ethical concerns in assessment (Liang et al., 2023), underscoring the need to integrate AI in ways that complement pedagogical principles and enhance, rather than replace, traditional instruction (Kartal and Yeşilyurt, 2024; Son et al., 2023).

2.4 NLP in Education

NLP, a subfield of AI, is pivotal in language education, powering chatbots and dialogue systems that provide interactive, meaning-focused practice often lacking in classrooms. Combined with Automatic Speech Recognition (ASR), NLP enables detailed speech analysis, error detection, and corrective feedback, improving fluency, pronunciation, and grammatical accuracy (Litman et al., 2018; Kang et al., 2024; Pérez-Ortiz et al., 2024). It also supports data-driven learning by retrieving relevant examples and generating exercises, enhancing personalized and actionable language practice (Bibauw et al., 2022; Koç and Savaş, 2024).

2.5 AI Tools for Speaking Development

Research demonstrates that AI and NLP tools effectively enhance L2 speaking skills across multiple domains, including pronunciation, fluency, and willingness to communicate (WTC). Tools providing phonetic-level feedback, such as NOVO and ASR-based CAPT systems, improve segmental accuracy, intonation, and stress (Bashori et al., 2024; Amrate and Tsai, 2024; Sun, 2023; Kang et al., 2024). Conversational AI and chatbots reduce speaking anxiety, boost confidence, and promote fluency and WTC (Du and Daniel, 2024; Ruan et al., 2021). Their interactive, game-like features increase engagement and motivation, with adaptive feedback further sustaining practice and progress (Dai and Wu, 2025; Kahlerras and Bennacer, 2025; López-Minotta et al., 2025). Beyond pronunciation, AI tools support vocabulary, grammar, and pragmatic competence, helping learners develop comprehensive linguistic skills (Xu and Cao, 2022; Ruan et al., 2021; Du and Daniel, 2024).

NLP and AI tools substantially enhance L2 speaking by providing safe, engaging, and personalized practice that reduces anxiety and builds confidence. They improve pronunciation, fluency, and willingness to communicate while boosting motivation through interactive feedback. Despite limitations like technological constraints and algorithmic bias, ongoing research points to more effective and equitable tools, with future progress relying on ethical design, interdisciplinary collaboration, and longitudinal studies.

3. Methodology

3.1 Research Design

This study adopted a quantitative, cross-sectional survey design to examine how NLP-based tools support speaking development among L2 learners. The design focused on measuring learner perceptions across multiple dimensions, such as fluency, pronunciation, confidence, feedback usefulness, and engagement. Because the study relied on self-reported Likert-scale responses, the analysis followed non-parametric statistical procedures, which were informed by data distribution checks.

3.2 Participants

A total of 61 individuals participated in the study, with a mean age of 25.13 years (SD = 4.72) and an age range of 18 to 37 years. The sample included 75.41% males, 21.31% females, and 1.64% who preferred not to disclose their gender. Regarding experience with NLP-based speaking tools, 77.05% of participants reported prior usage, while 21.31% had no experience. The frequency of tool usage varied, with responses ranging from never (6.6%), rarely (14.8%), sometimes (29.5%), often (21.3%), to very often (26.2%), reflecting a wide spectrum of engagement with NLP tools such as ChatGPT, ELSA Speak, and Duolingo AI.

3.3 Instrument

Data were collected using a structured online questionnaire with 10 Likert-scale items rated from 1 (Strongly Disagree) to 5 (Strongly Agree) via [Google Forms / Microsoft Forms]. The questionnaire measured key constructs including perceived improvements in speaking fluency, pronunciation accuracy, and confidence; the effectiveness of automated feedback; engagement and anxiety reduction; perceived academic value of integrating NLP tools into formal English programs; and perceived limitations such as speech misrecognition and technical issues. Composite scores were calculated for the main constructs to facilitate analysis.

3.4 Data Preparation and Normality Assessment

All survey responses were imported and processed in Python using Pandas and NumPy. Data cleaning included standardizing variable names, converting Likert-scale responses to numeric values (1–5), and excluding incomplete responses. Descriptive statistics such as median, interquartile range (IQR), mean, and standard deviation were calculated programmatically to ensure reproducibility. Normality of all ten items was assessed using the Shapiro–Wilk test, revealing significant non-normality ($p < 0.05$) and skewed, discrete distributions. Accordingly, all inferential analyses employed non-parametric methods appropriate for ordinal, non-normally distributed data.

3.5 Statistical, Correlation, and Reliability Analysis

Given the non-normal distribution of the Likert-scale items, non-parametric tests were applied for all group comparisons. The Mann–Whitney U test assessed differences between two independent groups (e.g., users vs. non-users of NLP tools, male vs. female participants), while the Kruskal–Wallis H test was used for variables with more than two groups, such as frequency of NLP tool usage. Significant Kruskal–Wallis results were followed by post-hoc pairwise comparisons using Dunn's test with Bonferroni correction. Ordinal relationships, including Age versus Benefits and Usage versus Confidence, were evaluated using Spearman's rank correlation coefficient (ρ). The questionnaire's internal consistency was assessed using Cronbach's alpha, with the Benefits Scale (7 items on fluency, pronunciation, confidence, feedback, engagement, and anxiety reduction) showing excellent reliability ($\alpha = 0.921$) and the Limitations Scale (2 items on speech misunderstandings and technical issues) showing acceptable reliability ($\alpha = 0.734$). These results confirm that both scales consistently measured participants' perceptions, supporting the validity of subsequent analyses.

4. Results

4.1 Descriptive Statistics and Response Patterns

4.1.1 Perceptions of NLP Tool Benefits and Limitations

Participants generally exhibited favorable perceptions of NLP tools across all evaluated dimensions. As illustrated in Table 1, participants reported positive perceptions regarding the benefits of NLP tools, with the highest mean scores observed for fluency enhancement (M = 3.98) and confidence building (M = 3.92). Table 2 reveals that limitations were also recognized, particularly concerning issues with speech recognition (M = 3.90). Despite these limitations, participants generally endorsed the integration of NLP tools into university programs (M = 3.70), as demonstrated in Table 3.

Table 1. Descriptive Statistics for Benefits Items (N = 61)

Item	Mean	SD	Median	IQR	Min	Max
Improved speaking fluency	3.98	0.93	4.00	1.00	1	5
Better pronunciation accuracy	3.88	0.85	4.00	0.00	1	5
Increased speaking confidence	3.92	0.85	4.00	0.00	1	5
Effective automated feedback	3.87	0.85	4.00	0.00	1	5
More engaging practice	3.48	0.99	4.00	0.00	1	5
Reduced speaking anxiety	3.70	1.18	4.00	1.25	1	5
Real-life speaking scenarios	3.75	1.13	4.00	1.00	1	5

Table 2. Descriptive Statistics for Limitations Items (N = 61)

Item	Mean	SD	Median	IQR	Min	Max
Speech/accent misunderstandings	3.90	0.97	4.00	1.25	1	5
Technical/internet issues	3.57	0.94	4.00	1.25	1	5

Table 3. Descriptive Statistics for Integration Recommendation (N = 61)

Item	Mean	SD	Median	IQR	Min	Max
University integration recommendation	3.70	0.91	4.00	1.00	1	5

4.1.2 Response Distributions

Upon examining the response distributions, it was noted that the majority of benefits demonstrated positively skewed distributions, with a significant concentration of responses at scores of 4 (Agree) and 5 (Strongly Agree). Conversely, the distributions for limitation items were more symmetric, indicating a balanced recognition of the tool's constraints. Approximately 77% of participants either agreed or strongly agreed that NLP tools enhanced speaking fluency, while 10% acknowledged occasional inaccuracies in speech recognition, and the remaining 13% maintained a neutral position (Figure 1).

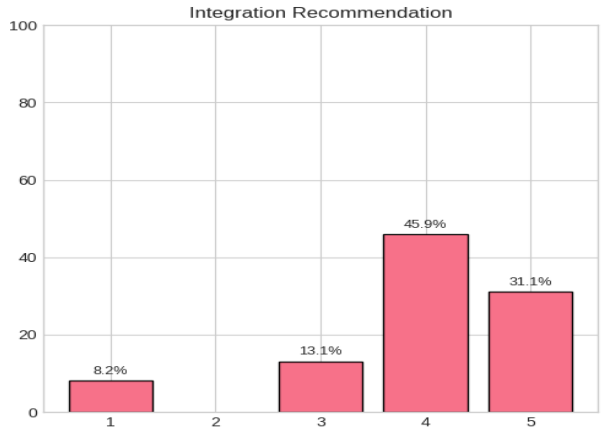

Figure 1. Integration Recommendation Distribution

4.2 Reliability of Measurement Scales

The internal consistency of the composite scales was evaluated using Cronbach's alpha. The 7-item Benefits Scale exhibited excellent reliability (α = 0.921), reflecting high internal consistency among items assessing fluency, pronunciation, confidence, feedback effectiveness, engagement, anxiety reduction, and scenario relevance. The 2-item Limitations Scale demonstrated acceptable reliability (α = 0.734) for the items addressing speech recognition issues and technical limitations.

4.3 Non-Parametric Group Comparisons

4.3.1 Users vs. Non-Users Comparison

The Mann-Whitney U test was conducted, where the U statistics are calculated as:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

where n_1 is the sample size of group 1, n_2 is the sample size of group 2, and R_1 is the sum of ranks for group 1. A statistically significant difference in perceived benefits between users of NLP tools (n = 47) and non-users (n = 13), U = 466.50, p = .0037. Users reported notably higher benefits (Mdn = 3.86, IQR = 0.71) in comparison to non-users (Mdn = 3.14, IQR = 0.71). The effect size was medium to large (Cliff's δ = 0.527), suggesting that approximately 52.7% of randomly selected scores from users would surpass those of non-users (Figure 2).

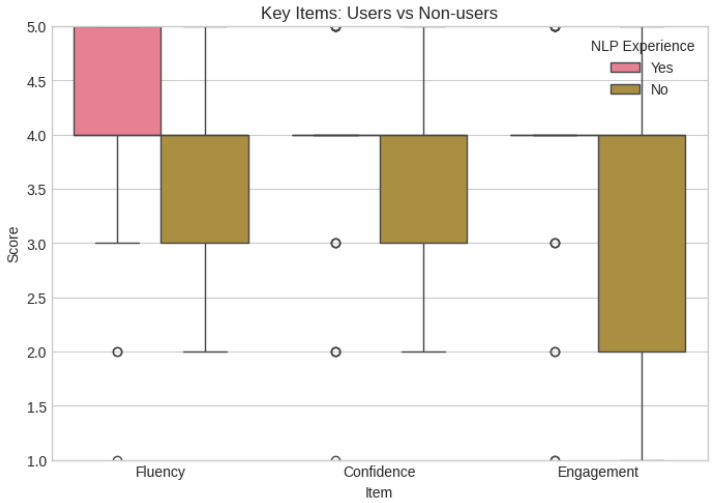

Figure 2. Key Items: Users vs Non-users

Figure 2. Key Items: Users vs Non-users, comparing the scores for fluency, confidence, and engagement between NLP tool users and non-users.

4.3.2 Gender Differences

No statistically significant differences were observed between male (n = 46) and female (n = 13) participants in terms of speaking confidence (U = 336.50, p = .4328) or engagement levels (U = 295.00, p = .7861). This indicates that the perceived benefits of NLP tools were consistent across gender groups.

4.3.3 Usage Frequency Effects

Participants were classified into three distinct usage categories: Low (Never/Rarely, n = 13), Medium (Sometimes, n = 18), and High (Often/Very Often, n = 29).

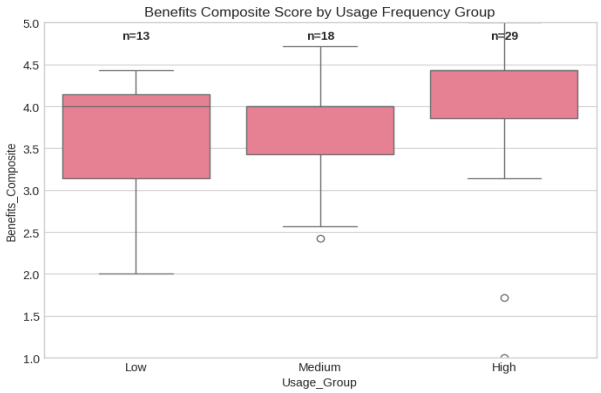

Figure 3. Benefits Composite Score by Usage Frequency Group

Figure 3 shows Benefits Composite Scores by Usage Frequency (Low, Medium, High). A Kruskal-Wallis test indicated a small, marginally significant difference, $H(2) = 6.10$, $p = .0473$, $\eta^2 = 0.071$. Post-hoc Dunn tests revealed no significant pairwise differences, though higher-frequency users tended to report greater perceived benefits (Mdn = 4.00, IQR = 0.71).

4.4 Correlation Analysis

4.4.1 Age and Perceived Benefits

Spearman's correlation analysis indicated no significant association between participant age and the perceived benefits of NLP tools (Figure 4),

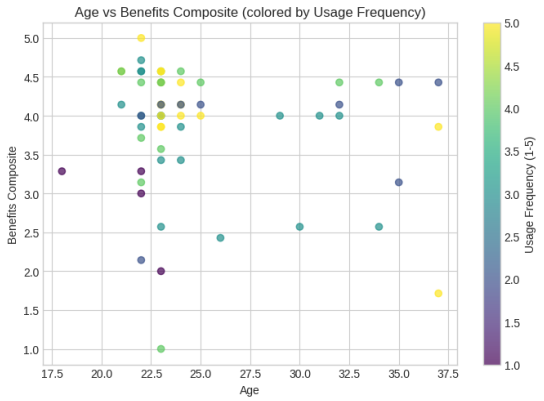

Figure 4. Age vs Benefits Composite (colored by Usage Frequency)

$\rho(59) = -0.124$, $p = .3459$. This suggests that perceptions of the tools were consistent across the age range examined.

4.4.2 Usage Frequency and Speaking Confidence

A significant positive correlation was found between usage frequency (ordinal scale) and speaking confidence, $\rho(59) = 0.442$, $p = .0004$ (Figure 5).

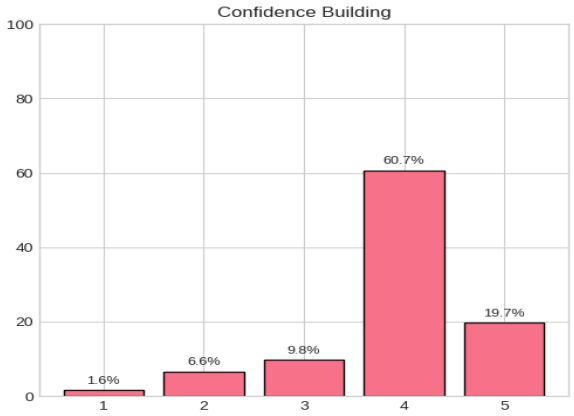

Figure 5. Confidence Building, showing the distribution of responses on confidence improvement through NLP tools.

This moderate correlation suggests that increased use of NLP tools is associated with greater self-reported speaking confidence.

4.4.3 Usage Frequency and Net Satisfaction

No significant correlation was identified between the frequency of usage and net satisfaction (calculated as benefits minus limitations), $\rho(59) = 0.182$, $p = .1634$. This suggests that although frequent users reported greater benefits, they also recognized limitations, leading to relatively consistent net satisfaction across different levels of usage.

4.5 Composite Scores Analysis

The distribution of composite scores (Figure X) revealed that participants generally perceived NLP tools as beneficial, with benefits (Mdn = 3.43, IQR = 0.71) outpacing limitations (Mdn = 3.75, IQR = 1.25). The net satisfaction score (benefits minus limitations), though close to zero (Mdn = -0.07, IQR = 1.07), indicates that while benefits and limitations are acknowledged, the overall positive perceptions of NLP tools suggest that the advantages are valued by most learners, despite some variations (Figure 6).

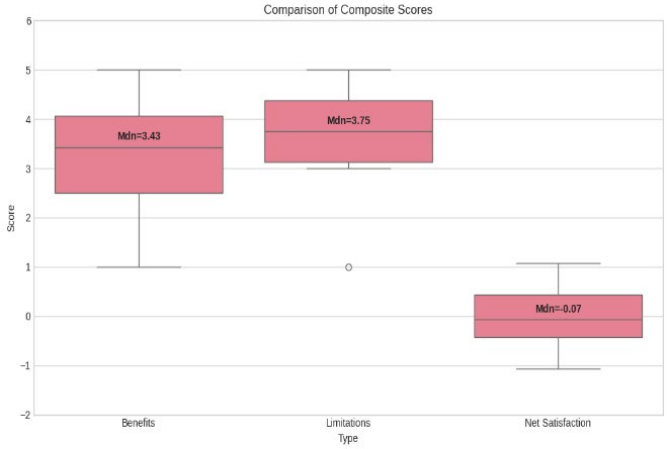

Figure 6. Comparison of Composite Scores

4.6 Key Findings Summary

Participants reported generally positive perceptions of NLP tools, noting improvements in fluency, pronunciation, confidence, and engagement (median ≈ 4.0). Users experienced greater benefits than non-users ($\delta = 0.527$, $p = .0037$), and more frequent use correlated moderately with confidence ($\rho = 0.442$, $p = .0004$), though usage-group differences were not statistically significant ($H = 6.10$, $p = .047$, $\eta^2 = 0.071$). While acknowledging limitations, perceptions were

consistent across gender and age, indicating that experience and frequency shape views of NLP tools' benefits while maintaining awareness of their constraints.

5. Discussion

5.1 Summary of Key Findings

This study investigated the perceptions of L2 learners regarding the benefits of NLP-based tools for speaking development. Utilizing a quantitative survey of 61 participants, the research found that learners generally hold positive views of NLP tools, particularly in enhancing speaking fluency, pronunciation accuracy, and confidence. The most significant finding was the clear disparity between users and non-users, with experienced users reporting substantially greater perceived benefits. Furthermore, a positive correlation was identified between the frequency of tool usage and self-reported speaking confidence. While limitations, especially concerning speech recognition accuracy, were acknowledged, participants overall supported the integration of these tools into formal language programs.

5.2 Research Question Findings and Literature Alignment

Learners reported that NLP tools significantly enhance speaking fluency ($M = 3.98$, $Mdn = 4.0$) and pronunciation accuracy ($M = 3.88$, $Mdn = 4.0$), with 77% agreeing or strongly agreeing that the tools improved fluency. Affective benefits included increased confidence (median = 4.0) and reduced anxiety (IQR = 1.25), and frequency of tool usage positively correlated with confidence ($\rho = 0.442$, $p = .0004$). Automated feedback was perceived as highly effective (median = 4.0), and users reported greater benefits than non-users ($U = 466.50$, $p = .0037$), highlighting the importance of active engagement. These findings align with prior research showing that NLP tools improve fluency, pronunciation, and confidence while reducing anxiety (Du & Daniel, 2024; Ruan et al., 2021; Bashori et al., 2024; Sun, 2023). Differences between users and non-users emphasize the role of engagement in learning (Li, 2024; Son et al., 2023), and frequent tool use reinforces confidence, promoting further practice (Kahlerras & Bennacer, 2025; López-Minotta et al., 2025). Participants also noted accent-related ASR limitations, confirming known challenges with L2 speech despite overall positive perceptions (Koç & Savaş, 2024; Amrate & Tsai, 2024).

5.3 Comparative Analysis of Current and Previous Research

This section presents a comparative checklist that highlights the strengths and unique aspects of this study in relation to previous literature. The comparison below illustrates how this research contributes to the growing body of knowledge on NLP tools for L2 speaking development (Table 4).

Table 4. Comparative Analysis of the Current Study and Previous Research on NLP Tools for L2 Speaking Development

Criteria	Current Study	Previous Studies
Sample Size	61 participants (quantitative survey)	Smaller sample sizes (typically 20-40 participants) in most studies
Focus of Study	Investigation of the benefits of NLP tools for L2 speaking development	Broader focus on technology's role in L2 acquisition, often not specific to speaking skills
Type of NLP Tools Analyzed	Multiple NLP tools (e.g., ChatGPT, Duolingo, ELSA Speak)	Focus on specific tools, such as ASR or CAPT, with limited comparison across tools
Experience of Participants	Comparison of users vs. non-users to assess differences in perceived benefits	Limited focus on user experience or prior familiarity with tools
Effect on Speaking Confidence	Positive correlation between usage frequency and self-reported speaking confidence	Confidence not directly measured in many studies
Type of Feedback Provided	Personalized, real-time feedback through NLP-based systems	Primarily general feedback, often lacking adaptation to individual needs
Acknowledgment of Limitations	Recognition of speech recognition accuracy issues, especially with accents	Focus on tool functionalities, less attention to limitations like speech recognition errors
Pedagogical Implications	Strong recommendation for integrating NLP tools into formal educational settings	General recommendations for technology use, without specific suggestions for integration into curricula
Innovative Contributions	Comprehensive evaluation of multiple NLP tools' impact across fluency, pronunciation, and confidence	Limited comparison of tools, typically focusing on a single aspect of language acquisition

5.4 Implications

This study provides empirical, learner-focused evidence on the role of NLP tools in language education, showing that interactions with these tools enhance not only speaking skills but also key affective variables such as confidence and anxiety. NLP tools are thus instrumental both for skill practice and for shaping a positive psychological environment that lowers the affective filter for L2 learners. The findings have practical implications for multiple stakeholders: educators and institutions are encouraged to integrate NLP tools into curricula via lab sessions, flipped classrooms, or blended assessments; developers should address issues like accent misunderstanding by diversifying training data with non-native speech (Kang et al., 2024; Li, 2024); and learners are validated in using these tools autonomously, with regular engagement linked to gains in perceived speaking skills, confidence, and overall language development.

5.5 Limitations of the Study

This study has several limitations. The small, male-skewed sample (N=61) limits generalizability, and reliance on self-reported Likert-scale responses may not accurately reflect actual speaking gains (Qiao & Yijun, 2023). Grouping diverse NLP tools together may obscure differences in their effectiveness and focusing solely on university-aged learners limits applicability to younger learners, professionals, or other cultural contexts. Additionally, the study did not examine how these tools were integrated pedagogically, restricting insights into best practices.

5.6 Recommendations for Future Research

Future research should use larger, more diverse samples and combine objective measures with learner perceptions to validate proficiency gains. Comparing different NLP tool types across varied cultural and professional contexts, and examining how tools are integrated pedagogically independently, in teacher-assigned tasks, or blended can clarify effectiveness and inform best practices.

6. Conclusion

This study adds a learner-perspective validation to the growing academic consensus on the value of NLP tools for L2 speaking development. It confirms that learners perceive these tools as effective allies in overcoming the classic hurdles of anxiety, limited practice, and lack of feedback. While mindful of the tools' current imperfections, particularly in speech recognition, learners see their overall benefit and advocate for their formal adoption in learning environments. The findings underscore that the future of language learning is not a choice between human instruction and AI, but a synergistic integration where technology handles scalable, personalized practice, freeing educators to focus on higher-order guidance, motivation, and the irreplaceably human aspects of communication. Future research must now focus on optimizing this partnership through better technology, informed pedagogy, and rigorous, longitudinal evaluation.

References

Amrate, M., and Tsai, P., Computer-assisted pronunciation training: A systematic review, ReCALL, vol. 37, no. 1, pp. 22–42, 2024.

Bashori, M., Van Hout, R., Strik, H., and Cucchiarini, C., I Can Speak: improving English pronunciation through automatic speech recognition-based language learning systems, *Innovation in Language Learning and Teaching*, vol. 18, no. 5, pp. 443–461, 2024.

Bibauw, S., Van Den Noortgate, W., François, T., and Desmet, P., Dialogue systems for language learning: A meta-analysis, *Language Learning & Technology*, vol. 26, no. 1, pp. 1–24, 2022.

Dai, L., and Wu, F., An AI-powered conversational system for college students learning English as a second language, *Education and Information Technologies*, vol. 30, no. 16, pp. 23393–23417, 2025.

Du, J., and Daniel, B. K., Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice, *Computers and Education Artificial Intelligence*, vol. 6, pp. 100230, 2024.

Dubey, P., Dubey, P., Raja, R., and Kshatri, S. S., Bridging language gaps: The role of NLP and speech recognition in oral English instruction, *MethodsX*, vol. 14, pp. 103359, 2025.

Jia, F., Sun, D., Ma, Q., and Looi, C., Developing an AI-Based learning system for L2 learners' authentic and ubiquitous learning in English language, *Sustainability*, vol. 14, no. 23, pp. 15527, 2022.

Kang, B. O., Jeon, H., and Lee, Y. K., AI-based language tutoring systems with end-to-end automatic speech recognition and proficiency evaluation, *ETRI Journal*, vol. 46, no. 1, pp. 48–58, 2024.

Kahlerras, W., and Bennacer, F., AI-driven approaches to advance speaking proficiency in LMOOCs: Insights, innovations, and pedagogical implications, *Journal of Studies in Language, Culture and Society (JSLCS),* vol. 8, no. 1, pp. 226–251, 2025.

Kartal, G., and Yeşilyurt, Y. E., A bibliometric analysis of artificial intelligence in L2 teaching and applied linguistics between 1995 and 2022, *ReCALL*, vol. 36, no. 3, pp. 359–375, 2024.

Kashinathan, S., and Aziz, A. A., ESL learners' challenges in speaking English in Malaysian classroom, *International Journal of Academic Research in Progressive Education and Development,* vol. 10, no. 2, 2021.

Koç, F. Ş., and Savaş, P., The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024, *ReCALL*, vol. 37, no. 1, pp. 4–21, 2024.

Li, R., Effects of mobile-assisted language learning on foreign language learners' speaking skill development, *Language Learning & Technology*, vol. 28, no. 1, pp. 1–26, 2024.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., and Zou, J., GPT detectors are biased against non-native English writers, *Patterns*, vol. 4, no. 7, pp. 100779, 2023.

Litman, D., Strik, H., and Lim, G. S., Speech Technologies and the Assessment of Second language speaking: Approaches, challenges, and opportunities, *Language Assessment Quarterly,* vol. 15, no. 3, pp. 294–309, 2018.

López-Minotta, K. L., Chiappe, A., and Mella-Norambuena, J., Implementation of artificial intelligence to improve English oral expression, *Multidisciplinary Journal of Educational Research,* vol. 15, no. 1, pp. 43–71, 2025.

Mora, J. C., Mora-Plaza, I., and Miranda, G. B., Speaking anxiety and task complexity effects on second language speech, *International Journal of Applied Linguistics*, vol. 34, no. 1, pp. 292–315, 2023.

Pérez-Ortiz, J. A., Esplà-Gomis, M., Sánchez-Cartagena, V. M., Sánchez-Martínez, F., Chernysh, R., Mora-Rodríguez, G., and Berezhnoy, L., A conversational intelligent tutoring system for improving English proficiency of Non-Native speakers via debriefing of online meeting transcriptions, *ACL Anthology,* 2024.

Qiao, W., and Yijun, C., The role of live transcripts in synchronous online L2 classrooms: Learning outcomes and learner perceptions, *Education and Information Technologies,* vol. 28, no. 11, pp. 14783–14804, 2023.

Ruan, S., Jiang, L., Xu, Q., Liu, Z., Davis, G. M., Brunskill, E., and Landay, J. A., EnglishBot: An AI-powered conversational system for second language learning, In *Proceedings of the 26th International Conference on Intelligent User Interfaces*, pp. 434–444, 2021.

Son, J., Ružić, N. K., and Philpott, A., Artificial intelligence technologies and applications for language learning and teaching, *Journal of China Computer-Assisted Language Learning*, vol. 5, no. 1, pp. 94–112, 2023.

Sun, W., The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation, *Frontiers in Psychology*, vol. 14, pp. 1210187, 2023.

Xu, F., and Cao, R., Epistemic stance in Chinese L2 Spoken English: The effect of Grade and Genre-Specific Questions, *Languages*, vol. 8, no. 1, pp. 15, 2022.

Appendix A: Survey Instrument

Section A: Demographic Information

1. Age: [Open response]
2. Gender: Male / Female / Other / Prefer not to say
3. Have you ever used NLP-based tools for speaking practice? (e.g., ChatGPT, ELSA Speak, Duolingo AI)
   Yes / No
4. How often do you use such tools?
   Never / Rarely / Sometimes / Often / Very Often

Section B: Perceptions Toward NLP Tools

(Scale: 1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly Agree)

1. NLP tools have improved my English-speaking fluency.
2. I have noticed better pronunciation accuracy since using NLP tools.
3. Using NLP tools has increased my confidence in speaking English.
4. The automated feedback from NLP tools helps me correct my mistakes effectively.
5. NLP tools make speaking practice more engaging and enjoyable.
6. I feel less anxious practicing English with AI tools than with human listeners.
7. NLP tools provide real-life speaking scenarios that match my learning needs.
8. Sometimes, NLP tools misunderstand my speech or accent, which affects feedback accuracy.
9. Technical or internet issues sometimes limit my ability to use these tools.
10. I believe NLP tools should be integrated into university English programs to enhance speaking development.

Biographies

S. M Mehedi Hasan is currently in his final year of undergraduate studies in Computer Science and Engineering at American International University-Bangladesh (AIUB), he focuses on areas such as embedded systems and computer vision, with a strong emphasis on technology-driven problem-solving. His future goals include conducting advanced research and developing practical applications in the field of intelligent systems.

Anika Tahmina Chowdhury is currently completing her final year of Computer Science and Engineering at American International University-Bangladesh (AIUB), her academic focus encompasses software development, human-computer interaction, and data-centric approaches. She visions advancing her education while working on technological innovations that improve accessibility and user experience.

Md. Shohaib Islam is currently in his final year of Computer Science and Engineering at American International University-Bangladesh (AIUB), he has academic interests span IoT systems, machine learning, and automation. He is eager to pursue innovative concepts and contribute to the development of technology.

Iftekhar Mahmud is an English lecturer at American International University-Bangladesh (AIUB) with a genuine commitment to student success. He holds a bachelor's degree in Linguistics and TEFL from AIUB and an MEd in TESOL from the University of Dundee. His English proficiency is evidenced by his IELTS band 8 scores, achieved twice in 2018 and 2021, reflecting his strong language skills and dedication to excellence in teaching.