https://doi.org/10.69639/arandu.v12i4.1780

# Using Grok to improve English speaking fluency in B1 English as a Foreign Language (EFL) Learners

*El Uso de Grok para mejorar la fluidez en inglés en estudiantes de inglés como lengua extranjera (EFL) de nivel B1*

**Franklin Absalon Mora Preciado**
franklinmora942@gmail.com
famorap@ube.edu.ec
https://orcid.org/0009-0008-6988-2845
Universidad Bolivariana del Ecuador
Durán-Ecuador

**Isabel Batista Medina**
isabelbm1956@gmail.com
https://orcid.org/0000-0002-1701-4957
Universidad de Oriente.Cuba

**Josué Reinaldo Bonilla Tenesaca**
jrbonillat@ube.edu.ec
https://orcid.org/0000-0002-6748-2345
Universidad Bolivariana del Ecuador
Durán-Ecuador

*Artículo recibido: 10octubre 2025 -Aceptado para publicación: 18 noviembre2025*
*Conflictos de intereses: Ninguno que declarar.*

## ABSTRACT

Many students of the B1 English level from the "Hispanoamericano" High School of Guayaquil-Ecuador have difficulties speaking in English with fluency due to fear of making pronunciation mistakes and because of lack of practice outside the school. This investigation was carried out with 50 B1 EFL learners, who attended classes on Saturdays. This research was based on an action research design supported with the Communicative Language Teaching approach. Additionally, Grok, an Artificial Intelligence (AI) app was used as the main tool to improve fluency by applying the Analysis- Design -Development -Implementation and Evaluation (ADDIE) Model. Furthermore, data was collected through the use of survey, analysis of chats generated with the voice input of students and the app. Moreover, the results showed students felt more relaxed and motivated when practicing speaking with the app and showed they made fewer pauses, and their answers were longer and more coherent. The survey indicated that the improvement in the trust and motivation average increased from 3.2 to 4.2. Finally, the conclusions and implications showed that the use of this app in a classroom improved fluency, lowered anxiety, and improved metacognitive strategies. Therefore, this research shows that the use of Artificial Intelligence tools

like Grok can be integrated in the teaching learning process of schools to improve English oral fluency levels.

*Keywords:* English-speaking fluency, Grok, B1 learners, communicative language teaching, metacognitive strategies

## RESUMEN

Muchos estudiantes del nivel B1 de inglés del Colegio "Hispanoamericano "de Guayaquil-Ecuador tienen dificultades hablando en inglés con fluidez debido al miedo de cometer errores de pronunciación y por la falta de práctica fuera de la escuela. Esta investigación fue llevada a cabo con 50 aprendices B1 de inglés como lengua extranjera que asisten a clases los sábados. La cual, está basada en un diseño de investigación-acción apoyado con el enfoque de Enseñanza Comunicativa de Lenguas. Además, Grok, una aplicación de Inteligencia Artificial (IA) fue usada como la herramienta principal para mejorar la fluidez aplicando el Modelo de Análisis, Diseño, Desarrollo, Implementación y Evaluación. Asimismo, los datos fueron recolectados a través del uso de encuesta, análisis de chats generados con la entrada de voz de estudiantes y de la aplicación. Igualmente, los resultados mostraron que los estudiantes se sintieron más cómodos y motivados cuando practicaban hablar con la aplicación, hicieron menos pausas con la herramienta, y sus respuestas fueron más largas y más coherentes. La encuesta indicó que la mejora en la media de confianza y motivación aumentó de 3.2 a 4.2. Finalmente, las conclusiones e implicaciones revelaron que el uso de la aplicación en un aula mejoró la fluidez, redujo la ansiedad, y optimizó el uso de estrategias metacognitivas. Por lo tanto, esta investigación refleja que el uso de herramientas de Inteligencia Artificial como Grok puede ser integrado al proceso de enseñanza aprendizaje de las escuelas para mejorar los niveles de fluidez oral en inglés.

*Palabras clave:* fluidez oral en inglés, Grok, estudiantes B1, enseñanza comunicativa de lenguas, estrategias metacognitivas

## INTRODUCTION

Fluent oral English is not just a skill; it is a ticket to the world. Yet countless EFL learners have difficulties when trying to speak confidently. Reaching a B1 level is the bridge between hesitation and global opportunity. Also, fluent English speaking opens global doors, but many EFL learners have a difficult time speaking fluently. If people wish to succeed in many areas in the world, they would need to be proficient, at least, in the B1 English-speaking level.

Examples of renowned authors are Krashen and Nation who have written about this matter. Krashen (1982), pointed out the importance of comprehensible input in language acquisition and stated that pronunciation and fluency come when learners practice meaningful tasks, instead of only being corrected. Also, Nation (2001) stressed the importance of vocabulary in fluency, signaling out that pronunciation and proficiency in speaking are made stronger with a large number of known words.

Research on specific AI apps, like Grok, is still limited. Liu et al. (2025) explained that AI tools, like ChatGPT, can help students learn English because they give feedback that helps them improve speaking, writing, and thinking skills. Wangsa et al. (2024) indicated that even though Grok is a relatively new AI tool, it is notable for its real-time interaction, and its features promise to help improve English fluency, such features resemble a real person with sarcasm and an excellent sense of humor, making the learning process very enjoyable. Thus, Grok seems to be a promising AI tool for helping EFL learners improve their English-speaking fluency.

It was observed a problem in students of third-year BGU of the “Hispanoamericano” High School of Guayaquil, Ecuador. The problem was that most of them could do all the macro skills in the B1 English level, but they could not speak, and the reasons why they could not interact orally were similar because they were nervous, insecure and felt embarrassed about mispronouncing words in front of their classmates and teachers. These problems made learners perform in English less proficiently in class and lose confidence in their speaking ability.

As Vygotsky (1934/1986) stated that speaking is what puts together thoughts and real communication, that without a dialogue, cognition stays in the brain and does not come out. Also, Dennis (2024) showed how AI conversational tools could give learners personal feedback to improve English fluency. Likewise, Shikun et al. (2024) explained that AI chatbots help students practice English speaking in a friendly manner, making learning more enjoyable without feeling nervous or judged. In this study, the independent variable was the use of Grok and the dependent variable was speaking fluency.

Thus, the main objective of this study is to demonstrate how this app, which is an AI-powered tool, can help with its real-time conversational modes and abilities to improve English-speaking fluency among third-year BGU B1 English level learners. This research aims to give a possible solution to the mentioned problem above by providing a safe, interactive environment.

Consequently, to evaluate if Grok app can address this issue, the study used a mixed-methods approach: a five-question Likert-scale, pre/post-tests to 50 third-year BGU students, who were selected intentionally. Surprisingly, they were officially B1 on paper, but mute in oral practice. The interventions consisted of six weekly interventions, each 25 minutes: two 10-minute Grok voice sessions with a five-minute break in between, applying Communicative Language Teaching (CLT), role-plays, and metacognitive self-checks to reflect on progress. Sessions were scheduled weekly to prevent cognitive overload.

In summary, this study looks for a possible solution to a problem of many B1-level EFL students: that is, those who can read, write, and listen in English, but they cannot communicate orally, due to nervousness about being embarrassed for making pronunciation mistakes in front of their peers and teachers. Thus, Grok, as an AI-power tool that can give real-time conversational practice, has the benefits that this app talks like a real person. In turn, learners do not feel nervous about making mistakes because they know that the app is not a person, who may criticize them if they distort words, hence, creating a safe and confident environment to improve English-speaking fluency.

On this respect, Mingyan et al. (2025) explained that AI mobile apps help students practice English-speaking skills since they give instant feedback and personalized guidance. Likewise, Quvanch et al. (2024) explained that anxiety makes many students panicky and lose concentration. Finally, the novelty of this study is about Grok app, specifically, a relatively new AI-powered application with interactive voice modes that resembles human conversations. Yet, this study was done only in one group of 50 B1 learners in a specific school, therefore, more research is needed to see the real impact of this app in different English-speaking levels and places.

**Study Rationale**

This study was inspired by Grok, a tool created by xAI, Elon Musk's company in 2023 which happens to give an interactive and enjoyable way to practice oral skills. As Mohammed et al. (2025) explained that new AI tools such as ChatGPT, DeepSeek, and Grok make learning English more creative and interesting, they also said that Grok is becoming a main player because it gives real-time, engaging support for students' communication practice.

Thus, many students in Guayaquil do not speak English well, possibly because they might be shy or may not practice enough outside classes. Similarly, they do not rehearse beyond the school classes, perhaps because they may not have the financial means to pay for private English classes. According to INEC (2023) About 62.2% of households have internet, and "72.7% of people aged 5+ used the internet in the last 12 months. So, Grok app is free, works on phones, and no fast internet is needed, so this app was a good choice for these students because it is a promising new AI tool with potential to help acquire the B1 English-speaking level.

For that reason, the researcher selected 50 students who could read, write, and listen in English, but could not speak. This moderately sized group allowed to analyze progress using an automated tool like Grok's transcription through voice input and teacher observations. No matter which voice mode was used, this tool knew a lot about many subjects and it could talk about any topic at any level, so it was a good pedagogical tool to help people speak English more fluently.

The goal was to verify if Grok could help students improve their English speaking, because many did not speak, perhaps because they get uneasy when they participate orally, and as a result, they use short sentences, and make pauses saying "um" or "eh", or they only answer with a "yes" or a "no" words. So here is where this tool comes in handy and can act as a friend, waiter, or teacher in voice conversations, so learners can practice real dialogs without feeling nervous.

This study adapts the methodological structure proposed by Aremu et al. (2025), who applies the ADDIE model, Vygotsky's ZPD, CLT, and Cognitive Load Theory in an Ecuadorian EFL context. However, the present research is different because it is done on high school B1 learners, and utilizes Grok, an AI conversational tool.

ADDIE was the model followed to plan carefully the study which consists of Analysis, Design, Development, Implementation, and Evaluation stages, and it was also included ideas from experts:

**Vygotsky (1978):** Students learn better with help from someone more knowledgeable, like Grok.

**Hymes (1972):** Speaking activities should be realistic, such as ordering food or asking for directions, among other communicative functions.

**Sweller (1988):** Tasks should be simple, manageable to avoid overload and confusion.

**Comparative reliability of voice modes: Grok vs. Chatgpt, Copilot, and Gemini**

The apps Grok, ChatGPT, Copilot, and Gemini were observed by the researcher for over six months, and based on the researcher's personal experience, the app that resembled a human being the most was Grok; also, it could keep a conversation for an unlimited period of time; additionally, the app felt as if it could reason like a human. On the other hand, ChatGPT was almost as good as Grok but it did not keep the conversation flowing; instead, the talk was stopped by the statement "if there is anything else I can do for you, just let me know." and finally, Copilot and Gemini felt like machines; they acted automated. Therefore, the app that resembled the most a human-like experience was Grok, thus, making the use of the tool more enjoyable.

**Theoretical framework**

The framework of this research was supported on prominent theories, like CLT, Vygotsky's ZPD, and Krashen's Input Hypothesis, to explain why Grok app promotes the acquisition of speaking English fluency and self-confidence.

**Communicative Language Teaching (CLT)**

Richards (2006) explained that CLT is all about using the language to be able to communicate in real-life situations, not just learning grammar rules. Also, as said by Hymes (1972), that communicative competence is equal with being capable of knowing when and how to use the language appropriately based on contexts. Additionally, with this method, students interact meaningfully in English and learn through oral communication instead of memorizing grammatical rules. Grok makes this possible through offering real conversation practice.

**Sociocultural Theory of Learning**

Vygotsky (1978, 1986) proposed in his theory that learning takes place through social interaction and use of tools and that guidance within the Zone of Proximal Development (ZPD) gives more productive learning. Because of this, Grok app goes a step further by acting as an electronic tutor, giving feedback, correction, and illustrations. It allows students to move from brief answers to more natural conversations. Also, it makes learners more confident and reduces fear when speaking in English.

**Input Hypothesis and Affective Filter**

On the other hand, Krashen (1982) proposed that students learn language as they understand messages just a little higher than their current level, and that anxiety or fear can stop this process by the affective filter. Also, they can make the attempt to speaking in a relaxed scenario with the aid of Grok. This is a proof that AI provides constructive criticism without judging, reducing the affective filter and making communication easier and more natural.

**Cognitive Load and Step-by-Step Learning**

Sweller (1988) argued that instruction should be planned in a way to avoid cognitive overload and advance from simple tasks to more difficult tasks progressively, in a scaffolding manner. Additionally, the ADDIE model used in this study follows this guideline, so it started with less complex activities and increased progressively to more difficult activities, and students were given immediate feedback that helped them stay focused and confident during learning.

**Artificial Intelligence in English Language Teaching**

Shikun et al. (2024) indicated that AI chatbots could help in language development by giving a low-stress place to practice. In this same line of reasoning, Liu et al. (2025) said that AI tools motivated learners to engage and have critical thinking when used in EFL contexts. Also, Dennis (2024) observed that AI speech recognition tools helped students with pronunciation and fluency. Therefore, artificial intelligence tools like Grok give learners more chances to practice English, receive instant feedback corrections, and learn at their own pace. Together with good teaching, AI is a better and promising tool to help improve oral fluency.

## METHOD

### Participants

The research group consisted of 50 students, all had the same sociocultural background, food, music, sports, and they were B1 English level, they could communicate simple ideas but made mistakes and paused often. For example, a student may say "I go shop" instead of "I'm going to the shop". They were chosen because the English teacher knew them from class, and they wanted to improve their speaking skills.

All participants were 18 and older, nevertheless, this research followed ethical rules to protect learners, their names and their voice generated by Grok chat texts were kept private (their real names were not used). Additionally, only the researcher saw the results, and the answers from the Google Forms surveys were stored safely online. These steps helped make sure that the study was honest, safe, and respectful for all students.

Table 1 below describes the population (which is the sample) before the interventions of this research.

**Table 1**

*Group Information*

| Strengths | Weaknesses | Opportunities | Threats |
|---|---|---|---|
| **Cohesive cohort of 50 adult learners (18+), all Spanish-speaking, easing Grok's code-switching (Table 3).** | No prior AI tool experience, increasing learning curve (Table 1). | Scalable to other schools, with 72.7% internet usage in Ecuador (INEC, 2023). | Unreliable internet or devices in low-resource settings (INEC, 2023). |
| **High motivation for university studies drives engagement (Table 1).** | Traditional, grammar-based instruction lacks CLT/CLIL integration (Table 1). | Grok's novelty invites further EFL research (Mohammed et al., 2025). | Competition from established AI tools (e.g., ChatGPT) (Comparative Reliability). |
| **Familiarity with mobile apps (WhatsApp, TikTok) supports Grok use (Table 1).** | Limited prior English exposure may slow progress (Table 1). | Potential curriculum adoption due to reduced anxiety and improved oral fluency (Table 7). | Privacy concerns regarding voice data storage (Ensuring Ethical Standards). |

### Sampling procedure

For carrying out this research, 50 students were chosen on purpose, because their English teacher stated that they needed help with speaking skill. Also, all of them lived in Guayaquil city and spoke Spanish at home. Additionally, many did not have fast internet nor good computers at home, so most of them used only their phones to interact with Grok. Additionally, their ages were 18 and older, however, parents' consent was asked to align with ethical standards and guidelines of the institution.

**Research design**

The interventions combined two strategies, metacognitive and spaced learning.

**Metacognitive Strategy**

A metacognitive strategy is a three-step cycle, which was used for planning, monitoring, and reflecting, to improve students' speaking skill. According to Meher et al. (2024), metacognition helps learners think about their own learning, and use strategies like thinking aloud, making idea- chart, and self-assessing to improve their performance in the school.

**Planning (before speaking):** Before each 10-minute Grok session, students wrote one personal goal on paper, for example: "I will speak in full sentences without saying 'um' or 'eh'."

**Monitoring (while speaking):** After five minutes, they stopped for ten seconds to ask themselves: "Am I reaching my goal? Am I too nervous?" Then, they made small changes and continued speaking.

**Reflection (after speaking):** After finishing, they answered three short questions:

How many pauses did I make?

Did I reach my goal?

What can I do better next time?

Table 2 below indicates the metacognitive steps with examples.

**Table 2**

*Metacognitive Stages*

| Stage | When | What Students Did | Example |
|---|---|---|---|
| **Planning** | Before speaking | Set a goal | "I will speak in full sentences." |
| **Monitoring** | During speaking | Checked progress | "Am I too nervous?" |
| **Reflection** | After speaking | Reviewed performance | "Next time I'll pause less." |

**Spaced Learning Strategy**

The spaced learning strategy is about studying or practicing in short parts with breaks in between, instead of studying in one long, single session without taking a break. As Kamali et al. (2024) explained that using spaced learning, with reviewed lessons and short breaks in between helps students remember information better and feel less nervous about learning. In EFL learning, it helps learners remember English and not feel tired or bored, also, the breaks in between short study sessions give the brain time to rest and store the new information, so one example of spaced learning is when one learner studies for ten minutes, rests for five minutes, and then reviews again for another ten minutes or when they practice English speaking with Grok to speak English in two ten-minute sessions, with a five-minute break in between, in total, the practice lasted twenty-five minutes. Also, participants are encouraged to do the same practices at home to make sure what has been learned stays in their memory.

The intervention consisted of six weekly interventions, each 25 minutes: two 10-minute Grok voice sessions with a five-minute break in between, conducted on Saturdays, applying Communicative Language Teaching (CLT) role-plays and metacognitive self-checks to reflect on progress. Then, after each intervention, students were encouraged to practice speaking at home using Grok until they felt comfortable, as a strategy to overcome nervousness of speaking in front of people. The researcher used three main ways to measure progress: surveys before and after the study, a tool to analyze their chat conversations.

The English teacher was asked to apply the Grok activities throughout all the six-week period with key CLT principles through meaningful tasks and students' interaction. Richards (2006) states that CLT encourages real-life communication and meaningful language use.

There were 12 different voice chats or topics chosen by the teacher, two per week, over a six-week intervention period, each with Grok acting in a different role, such as a waiter, friend, or travel agent. The app also has different voice modes that can be applied to each role. The interventions took place in the virtual classroom. The teacher used his phone as a microphone, then called on learners one by one to ask questions to the app according to the tasks from Table 3 below. Each intervention lasted 10 minutes with a five-minute recess in between. All participations were saved by the tool in written form, which answered and corrected mistakes, for example, changing "I want pizza" to "I'd like pizza." Students also practiced at home, and afterward tips were sent via WhatsApp, such as "Try using longer sentences."

Table 3 below indicates the use of CAP through 12 voice chats activities for 50 students through

interaction with Grok.

**Table 3**

*Plan for role plays using CAP tool*

| Talk Number | Topic | Grok's Role | Student's Task | What We Checked through CAP |
|---|---|---|---|---|
| 1 | Greetings | Friend | Say who you are | Pauses, words, logic |
| 2 | Travel | Travel agent | Book a trip | Pauses, words, logic |
| 3 | Food | Waiter | Order food | Pauses, words, logic |
| 4 | Family | Cousin | Talk about your family | Pauses, words, logic |
| 5 | Shopping | Store worker | Ask for clothes | Pauses, words, logic |
| 6 | Weather | News reporter | Talk about weather | Pauses, words, logic |
| 7 | Home | Neighbor | Describe your house | Pauses, words, logic |
| 8 | School | Teacher | Talk about school | Pauses, words, logic |
| 9 | Job Interview | Boss | Answer job questions | Pauses, words, logic |

| Talk Number | Topic | Grok's Role | Student's Task | What We Checked through CAP |
|---|---|---|---|---|
| 10 | Hobbies | Friend | Share what you like | Pauses, words, logic |
| 11 | Daily Routine | Coworker | Describe your day | Pauses, words, logic |
| 12 | Review | Counselor | Say what you learned | Pauses, words, logic |

The researcher used the ADDIE model to plan the study that stands for Analysis, Design, Development, Implementation, and Evaluation stages.

**Analysis**

The researcher first talked with students using Grok to understand why they did not speak English well, so learners said they were afraid of making mistakes and did not practice outside class, then the researcher also checked school records and saw that less than half of them did not have a computer at home, so the app was a good choice because it could work on phones and did not need fast internet and it was user friendly.

**Design**

The researcher created 12 B1 level talks topics, such as food, travel, or family, and, each talk was short so that students would not get bored, then Grok asked easy questions at first and not too easy questions later on, to help them learn step by step.

**Development**

The tasks were elaborated in a way that they were not too difficult for students. For example, "Do you like pizza?" is easier to answer than "Explain why do you like pizza?" or it is easier to ask "Do you have a brother and a sister?" than to ask "Do you have any siblings?" because the word siblings is not too common for most beginners, then the app was used to act as a partner in the role play activities, at the same time, as the tasks developed, the app also corrected mistakes by saying "perhaps you meant to say this", for example, if a student said "they does", the app would correct the mistake by saying "perhaps you meant to say 'they do'".

**Implementation**

The implementation was a straightforward process, the teacher would connect his phone to a speaker, which would serve as a microphone, also the phone would have the app ready in the conversation mode, so at this point the app could listen and answer to anybody speaking in the classroom, and so the teacher would point out each student to participate by asking a question to the app according to the topics from Table 3, then the app would answer what students asked and also would give feedback.

**Evaluation**

The researcher analyzed surveys and chat data. The same survey was applied at the beginning of the six-week study and at the end so a comparison of the results of the survey were made. Also, the voice chats in written form were analyzed using CAP.

The ADDIE model helped the researcher plan the study in a clear and a step by step organized way, from understanding where students' English knowledge was at the beginning of the interventions, to checking their progress, to the final results, to aid the researcher use the app better to help improve fluency. This model facilitated understanding how students were more willing to speak and take part in English conversations. And also, the app's features like corrections, practice, and easy design helped them keep learning. Finally, the process showed the tool can be useful and motivating to improve speaking fluency in B1-level learners.

**Instruments**

The researcher used two main tools to see if students improved their English speaking:

**Survey (Before and After)**

Students were given a short survey with five statements to understand how they felt about speaking English using Grok. The statements used a one to five scale, where one is "I totally disagree" and five is "I totally agree." The questions were:

I feel good using Grok to practice English.

I want to speak English more with Grok.

I feel confident when I speak English.

I use correct grammar when I talk with Grok.

Talking with Grok helps me use English in real life.

The same survey was sent through Google Forms, and all 50 students answered both before and after the study.

**Conversation Analysis Protocol (CAP)**

This tool analyzed how learners spoke in Grok chats. The researcher created the CAP tool based on voice chats. It was measured three things:

Pauses: How often students say "um" or "eh."

Turn Length: How many words they use each time they answer.

Logic Level:

Level 1: Just answer the question.

Level 2: Answer and ask a question back.

Level 3: Start new ideas.

Table 4 below describes the voice modes of Grok.

**Table 4**

*Grok's voice modes*

| Voice Mode | Description | Platform Availability | Key Strengths for EFL Learners | Potential Drawbacks |
|---|---|---|---|---|
| **Assistant** | General-purpose mode with a neutral American accent; handles broad topics. | Computer (6 total), Mobile (13 total) | Versatile for everyday practice; covers all topics at any CEFR level. | None notable—safe and reliable. |
| **Unhinged Comedian** | Upbeat, humorous tone; excels at jokes but may use profanity. | Both | Builds fluency through fun, casual dialogue; encourages expressive speaking. | Profanity may not suit formal/school settings. |
| **Grok 'Doc'** | Doctor-like person; provides health/medical advice with deep knowledge. | Both | Useful for specialized vocabulary (e.g., medical terms); simulates real consultations. | Advice is informational, not a substitute for professionals. |
| **Argumentative** | Debates topics, provides reasoned arguments, and encourages user input. | Both | Promotes critical thinking and debate skills; great for advanced levels (B2+). | Can feel confrontational for beginners. |
| **(Other Modes)** | Additional modes (e.g., up to 13 on mobile) include variations like storyteller or expert roles (not detailed here). | Mobile-exclusive for extras | Expands role-playing for diverse scenarios, improving adaptability. | Availability may vary by update. |

Table 5 below describes the voice tones of Grok.

**Table 5**

*Grok's voice tones*

| Tone Category | Options | Description | Benefits for EFL |
|---|---|---|---|
| **Female** | Two variations (e.g., warm/neutral) | Softer, engaging tones that can adapt to conversational flow. | Encourages comfort for shyer learners; aids pronunciation practice. |
| **Male** | Two variations (e.g., authoritative/casual) | Deeper, varied pitches for a broader range of simulated interactions. | Builds confidence in diverse dialogues; mimics real-world variety. |

## RESULTS

### Data Storage Approach

The researcher saved two types of data:

Surveys: Students answered surveys on Google Forms, the results were saved on Google Sheets. Students answered the same survey at the beginning of the study and at the end.

Voice Chats in written form: Grok saved 600 voice chats in written form (50 students × 12 talks), these chats were later pasted on a word document. Each chat included:

What the student said

Grok's correction

Metrics: pauses, words, and logic level

Example:

Student Ana did Talk 3 from Table 3: "I want pizza, uh, please."

Grok: "Say 'I would like a pizza, please.'"

Metrics: 1 pause, 4 words, Logic Level 1

To demonstrate individual progress, see Table 6 below: Ana and Juan's voice chats before and after Grok intervention.

**Table 6**

*Example Chat stored*

| Student | Talk | Student's Task | Student's Sentence | Grok's Corrections | Pauses | Words | Logic |
|---|---|---|---|---|---|---|---|
| **Ana** | 3 | Order food | I want pizza, uh, please | Say, "I would like pizza, please." | 1 | 4 | 1 |
| **Ana** | 12 | Say what you learned | I learned to speak better without ums | None needed | 0 | 8 | 2 |
| **Juan** | 6 | Talk weather | It rain, um, bad | Say, "It is raining." | 1 | 4 | 1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Juan** | 12 | Say what you learned | I talk more now, what is your job? | None needed | 0 | 8 | 2 |

**Data Analysis**

The researcher examined the data in simple ways to see if students improved.

In surveys, the average (mean) was calculated, the median, the mode which is the most common answer, and the standard deviation to see how much the answers varied. For example, students' confidence started with a mean of 3.2 and increased to 4.2 after using Grok.

On the voice chats, which are the recorded voice inputs, pauses like "um" were counted, number of words per answer, and logic level.

The focus centered on showing changes in students' speaking, chats were read to find improvements, for example, "I am going" instead of "I going". And the survey was analyzed to see the level of comfort they had in speaking after using the app as shown in Table 7 below.

This table indicates the survey results, calculating the mean, median, mode, and standard deviation before and after the use of Grok.

**Table 7**

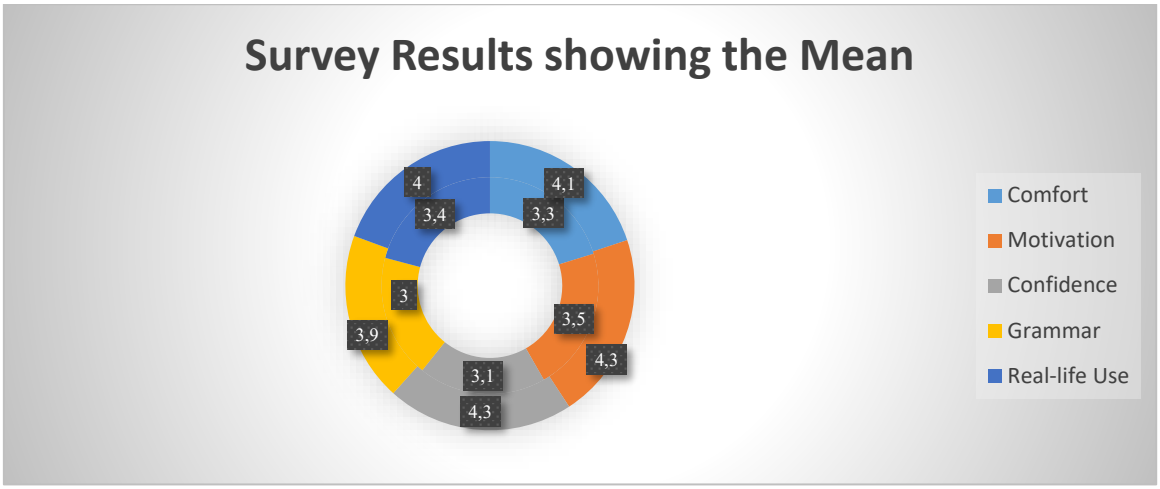
*Survey Results*

| Question | Before Mean | Before Median | Before Mode | Before SD | After Mean | After Median | After Mode | After SD | Change |
|---|---|---|---|---|---|---|---|---|---|
| **Comfort with Grok** | 3.3 | 3 | 3 | 0.8 | 4.1 | 4 | 4 | 0.9 | +0.8 |
| **Motivation** | 3.5 | 4 | 4 | 0.7 | 4.3 | 4 | 5 | 0.8 | +0.8 |
| **Confidence** | 3.1 | 3 | 3 | 0.9 | 4.3 | 4 | 4 | 0.8 | +1.0 |
| **Grammar** | 3.0 | 3 | 3 | 0.8 | 3.9 | 4 | 4 | 0.7 | +0.9 |
| **Real-Life Use** | 3.4 | 3 | 3 | 0.9 | 4.0 | 4 | 4 | 0.8 | +0.6 |

Chart 1 below shows the mean scores from Table 7.

**Chart 1**

*Survey Changes*




**Chat analysis results**

The results from the chats show that there was an overall fluency improvement, during the first interventions, students used short sentences with mistakes and many pauses like “I go uh shopping” instead of “I am going shopping” and towards the end of the intervention period, they used more words per answers with less pauses, and some of them showed logic level 2 according to the use of CAP analysis as shown in Table 8. Grok helped by correcting mistakes and giving feedback.

Table 8 below shows the chat results through the use of CAP.

**Table 8**

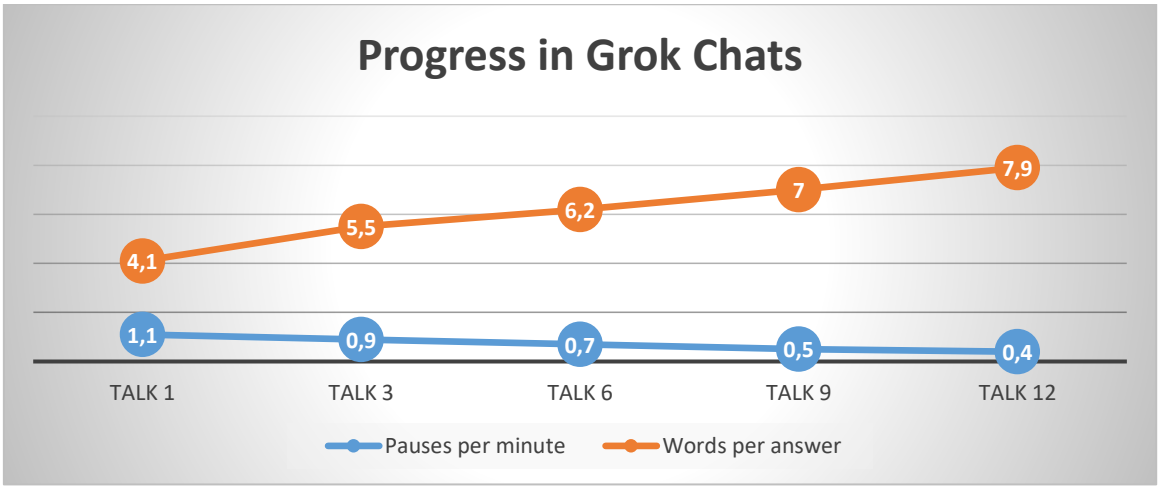
*Chat Results using CAP*

| Talk Number | Pauses (Mean) | Words per Answer (Mean) | Logic Level (Mean) |
|---|---|---|---|
| 1 | 1.2 | 4.2 | 1.0 |
| 3 | 0.9 | 5.5 | 1.1 |
| 6 | 0.6 | 6.3 | 1.3 |
| 9 | 0.5 | 7.0 | 1.5 |
| 12 | 0.4 | 7.8 | 1.8 |

Chart 2 below shows progress from voice chat converted to text by Grok.

**Chart 2**

*Chat Progress*




**Study Observations**

The researcher checked each student's chats carefully as shown in Table 9 below, Maria had six pauses in Talk one in the first week but only one pause in Talk 12 on the sixth week.

Carlos used three words per answer first week, but eight words by the last week, and, he even asked Grok an additional question: "What is your job?", showing Level 2 logic according to CAP.

Table 9 below indicates the results for 10 students from the 50-student population using CAP.

**Table 9**

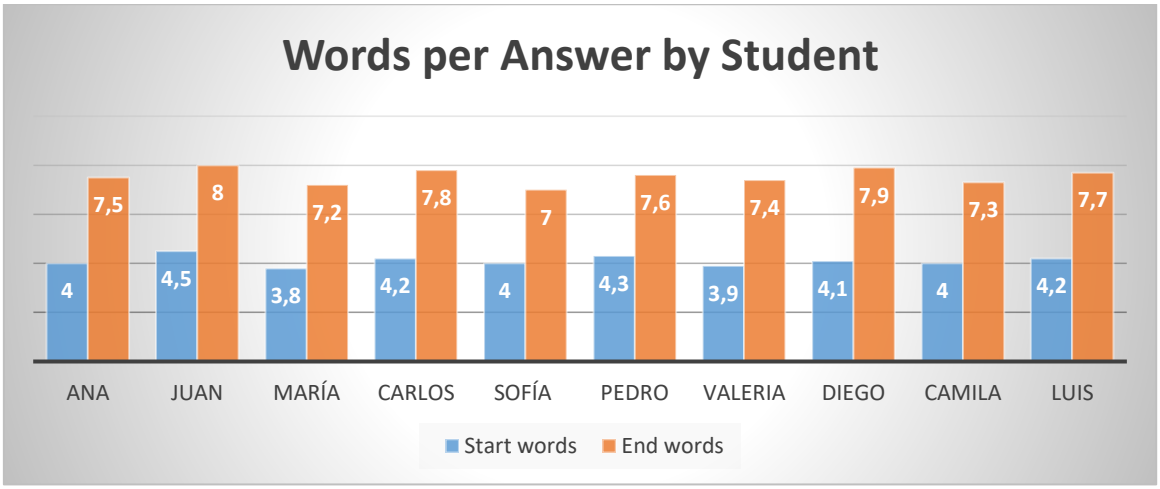
*Group Progress*

| Student | Pauses Week One | Pauses Week Six | Words Week One | Words Week Six | Logic Week One | Logic Week Six |
|---|---|---|---|---|---|---|
| Ana | 1.2 | 0.4 | 4.0 | 7.5 | 1.0 | 1.8 |
| Juan | 1.3 | 0.5 | 4.5 | 8.0 | 1.0 | 2.0 |
| Maria | 1.5 | 0.3 | 3.8 | 7.2 | 1.0 | 1.7 |
| Carlos | 1.1 | 0.4 | 4.2 | 7.8 | 1.0 | 2.0 |
| Sofia | 1.0 | 0.3 | 4.0 | 7.0 | 1.0 | 1.5 |
| Pedro | 1.4 | 0.5 | 4.3 | 7.6 | 1.0 | 1.8 |
| Valeria | 1.2 | 0.4 | 3.9 | 7.4 | 1.0 | 1.6 |
| Diego | 1.3 | 0.3 | 4.1 | 7.9 | 1.0 | 1.9 |
| Camila | 1.1 | 0.4 | 4.0 | 7.3 | 1.0 | 1.7 |
| Luis | 1.2 | 0.5 | 4.2 | 7.7 | 1.0 | 1.8 |

Chart 3 below shows examples of individual word count per answer, from week one chat blue to week six chat orange.

**Chart 3**

*Group Progress*




**Study limitations**

The population-sample for this research was only 50 students, so more study needs to be done on this.

Grok saves the voice input into written form, so the results from home practices could not be heard, they could only be read.

**Ensuring ethical standards**

Parents were asked to sign consent forms for students, but they could leave the study at any time. We used fake names, like Ana or Juan, to protect privacy. All chats were kept on a locked phone and only the researcher and the English teacher had access to it, and all procedures complied with the ethical guidelines and regulations of the "Hispanoamericano" High School and the Universidad Bolivariana del Ecuador to make sure the protection and confidentiality of student information.

**Grok's Effectiveness**

Grok was helpful because it is fast, free, funny, easy to use and learners liked talking to it because it is not a real person, it is a "robot", so they did not feel judged or embarrassed of making pronunciation mistakes as they would in front of real people.

Surveys showed that 80% of students felt comfortable using the tool, and 86% felt more confident when speaking English.

Chats showed that students used more words and paused less over time.

## DISCUSSION

Grok has many features, but the ones that are of interest for this research are the conversational ones, and it helps to point out that these features are almost identical, either in a computer or in a mobile phone. The only difference is that in a phone, the app has more "voice modes (roles)", but in spite of this, the app is the same. So, Grok's conversational features are

basically three, namely, the voice modes, the names of the voices that can resemble either a man's or a woman's voice, and the ability to imply and infer meaning.

So, Grok's roles act like real life impersonations, for example, in a computer, the tool has six roles, but in a phone, it has thirteen. For instance, one Grok's role is "Assistant", this role is very general in the way it talks, and, most of the time, it has a normal American accent, and it can answer all types of general topics. Besides this, there is also a role called "Unhinged Comedian" which has a different tone from the "Assistant", the "Unhinged Comedian" has an upbeat tone of voice, and it is excellent being funny and telling jokes, however, it might not be for everybody because it can talk profanity.

Additionally, another voice mode that is worth mentioning is "Grok 'Doc'", this acts as a doctor, and, amazingly, it can give recommendations about health and medicine, it actually knows a great deal about the medical field. Another role is the "Argumentative" one, it is interesting in the sense that it can actually argue about any topic, and gives the basis for its points of view, and it actually encourages people. Furthermore, regardless of the role used, Grok has an almost endless amount of knowledge of any field into it, so it can talk about any topic in any level so makes it a promising tool to improve fluency in the English-speaking skill.

The second Grok's voice feature is that the user can choose from four different tones of voice, it has two female, and two male tones, this feature makes it versatile and encouraging.

And the third conversational feature, and perhaps the most important one, is the ability Grok has to imply and infer meaning, it is almost as if Grok can reason like a human would, this feature enables it to maintain a human like conversation, it listens and responds, and keeps the subject going for as long as a person talks to it, and if an EFL learner asks to talk slower, or in any of the six English levels A1, A2, B1, B2, C1, or C2 according to the Common European Framework of Reference for the Languages (CEFR), it would.

For the aforementioned reasons: "voice modes", "tones of voice", and "inference, implying, and almost reasoning" are what make it a promising AI tool that can be implemented in schools to help students progress from one English speaking level to the next by imitating original and authentic settings that may support oral fluency.

After using Grok, students improved their English speaking in many ways. They spoke longer sentences and paused less. At the beginning, they produced short sentences like "I go shop" and paused about 1.2 per answer, and by the last talk, learners used around 7.8 words per answer and paused only 0.4 times.

Some students started asking questions back, like "Why is it raining?" showing that they could use Level 2 logic. Also, common mistakes, such as saying "I go" instead of "I'm going" or "I like pizza" instead of "I would like a pizza," were corrected immediately by the app and learners tried again and improved.

Survey results also showed improvement. Confidence increased from 3.1 to 4.2, Comfort from 3.3 to 4.1, and Motivation from 3.5 to 4.3. Additionally, students said they felt less anxious using the app. For example, Ana said, “Grok is like a friend, I feel good using Grok,” and Juan added, “I talk more now, no worry.” These results show that it helped them speak more freely, make fewer mistakes, and gain confidence.

Additionally, overall, the ADDIE plan worked: starting with easy questions, gradually increasing difficulty, and using role-play and immediate corrections helped students improve, also, the combination of short, repeated sessions, feedback, and student`s reflections using metacognitive strategy made learning safe and effective.

According to the general objective of analyzing the impact of Grok on learners’ academic performance, this study demonstrated that the use of this AI tool can be a possible solution to improve students’ English speaking fluency, with more participation and confidence. Besides, the data collected through CAP and surveys indicated that it is a valuable tool for helping improve the teaching and learning process in English as a foreign language in general Education.

## CONCLUSION

In conclusion, this study showed the effectiveness of the proposed solution, which is the integration of Grok app that significantly improved English-speaking fluency. It also, reduced communication anxiety, and increased confidence in B1 students of “Hispanoamericano” High School in Guayaquil-Ecuador. Additionally, scores for confidence improved from 3.1 to 4.2, also, the average number of pauses per answer decreased from 1.2 to 0.4, and response length increased from 4.2 to 7.8 words. Based on Communicative Language Teaching (CLT) and metacognitive strategies, and the ADDIE model, the intervention gave students a safe place to talk without fear of being judged that helped them grow English language fluency reaching the objectives. Although the population-sample was only of fifty students, the results showed strong proof that Grok is a helpful tool that can be used in any classroom to improve English fluency. Therefore, this study suggests that Grok can potentially be used in classrooms and at home as a complimentary help to improve English fluency.

## REFERENCES


Aremu, I. O., Paredes Espinosa, K. E., Intriago Cañizares, F., & Bonilla Tenesaca, J. R. (2025). AI-powered podcast interventions for enhancing speaking skills in English language teaching (ELT) adult A1 students. Arandu Utic, 12(3), 2702–2723. https://doi.org/10.69639/arandu.v12i3.1509

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.

Dennis, N. K. (2024). Using AI-powered speech recognition technology to improve English pronunciation and speaking skills. IAFOR Journal of Education: Technology in Education, 12(2), 107–126. https://doi.org/10.22492/ije.12.2.05

Hymes, D. (1972). Sociolinguistics: Selected readings (J. B. Pride, Ed.). Penguin.

Instituto Nacional de Estadística y Censos. (2023). Tecnologías de la información y comunicación (tic) en los hogares. https://www.ecuadorencifras.gob.ec

Kamali, M., Singh, M., Pallathadka, H., Etkov, I., Prasad, K. D., Ismail, S. M., … & Adefila, A. (2024). The efficacy of the spaced learning teaching method on student engagement and learning anxiety. Research Square (Preprint). https://doi.org/10.21203/rs.3.rs-4712072/v1

Krashen, S. D. (1982). Principles and practice in second language acquisition. Pergamon Press.

Liu, J., Sihes, A. J., & Lu, Y. (2025). How do generative artificial intelligence (AI) tools and large language models (LLMs) influence language learners' critical thinking in EFL education? A systematic review. Smart Learning Environments, 12(48), 1–35. https://doi.org/10.1186/s40561-025-00406-0

Meher, V., Baral, R., & Bhuyan, S. (2024). Assessing the impact of metacognitive interventions on academic achievement of higher secondary school students in education. Indonesian Educational Research Journal, 1(3), 144–157. https://doi.org/10.56773/ierj.v1i3.27

Mingyan, M. N. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: An empirical investigation. Humanities and Social Sciences Communications, 12(370), 1–15. https://doi.org/10.1057/s41599-025-04688-0

Mohammed, A. A. Q., Mudhsh, B. A., Bin-Hady, W. R. A., & Al-Tamimi, A. S. (2025). DeepSeek and Grok in the spotlight after ChatGPT in English education: A review study. Journal of English Studies in Arabia Felix, 4(1), 13–22. https://doi.org/10.56540/jesaf.v4i1.114

Nation, I. S. (2001). Learning vocabulary in another language. Cambridge University Press.

Quvanch, Z. Q. (2024). Analyzing levels, factors, and coping strategies of speaking anxiety among EFL undergraduates in Afghanistan. Cogent Education, 11(1), Article 2413225. https://doi.org/10.1080/2331186X.2024.2413225

Richards, J. C. (2006). Communicative language teaching today. Cambridge University Press.

Shikun, S., Grigoryan, G., Huichun, N., & Harutyunyan, H. (2024). AI chatbots: Developing English language proficiency in the EFL classroom. Arab World English Journal, 15(3), 292–305. https://doi.org/10.24093/awej/ChatGPT.20

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Vygotsky, L. S. (1986). Thought and language (A. Kozulin, Ed. & Trans.; Original work published 1934). MIT Press.

Wangsa, K. K. (2024). A systematic review and comprehensive analysis of pioneering AI chatbot models from education. Future Internet, 16(7), 219. https://doi.org/10.3390/fi16070219

## ANNEXES

APPENDIX A: Solicitud de aprobación presentada a la Universidad Bolivariana del Ecuador para el trabajo de titulación. https://drive.google.com/file/d/1NGlWuf7X4gn5ROLNsnbw48cArxHzCirV/view?usp=sharing

APPENDIX B: (i) Solicitud de autorización enviada al rector de la Unidad Educativa Hispanoamericano. (ii) Autorización otorgada por el Economista Felipe Bocca, rector. https://drive.google.com/file/d/18MqBiNfAjdA0g4G0vfi0Tt2BYa9rxgRW/view?usp=sharing

APPENDIX C: Encuesta Likert pre- y post-intervención. https://drive.google.com/file/d/1gAFN2HTGtnyEHta1G13eddQv8BgZpZZ_/view?usp=sharing

APPENDIX D: Protocolo de Análisis de Conversación (CAP) https://drive.google.com/file/d/1bKlD5ds8jNajuHIIjgR4ZilcMTtGZQZb/view?usp=sharing