AI for speaking skills assessment in foreign language acquisition

Olha Yanholenko[1,†], Antonina Badan[1,†], Nunu Akopiants [1,†], and Nataliia Onishchenko[2,†,*]

[1] National Technical University "Kharkiv Polytechnic Institute", Kyrpychova str. 2, Kharkiv, 61002, Ukraine

[2] Vasyl Karazin National University Kharkiv, 4, Svobody Sq, Kharkiv, 61022 Ukraine

Abstract

The present study investigates the benefits and flaws of speaking and pronunciation assessment in foreign language acquisition by both AI - driven technologies and their competitive counterparts, human experts. The experiment design rests on the comparison and efficacy analysis of the two opposing means of speaking assessment: those by AI tools and the more traditional human expertise. The conclusions drawn are meant to fill in the gap of AI platforms development best suited in their accuracy to fit into traditional learning based on immersion and simulations as prerequisites of AI tools integration with conventional educational methods. As a result, the theory of human-computer interaction is supplemented with new ideas of ecological approach for further improvement of the forthcoming AI advances in the present-day integration of digital and human methods of enhancing the prospects of foreign language speaking proficiency on the background of innovative findings in the domain of eco-linguistics.

Keywords

AI technologies, speaking assessment platforms, simulation, immersion, human-computer interaction, blending, ecolinguistics

1. Introduction

The present-day traditional learning and teaching designs in foreign language acquisition have long become dependent on the rapidly developing realm of AI technologies since the 2020s. Their prerequisites in ESL were multimedia platforms based on the idea of creating artificial foreign language environments for education by means of simulation [5, 7, 12, 13], learners immersion in a language environment [4, 5, 6, 12, 14], as well as the newly arising blending in teaching methods [1, 2, 3, 4, 5, 7, 8, 10, 11, 12, 13, 14, 16].

Thus, whereas the key requirements in foreign language learning, those of simulation, immersion and integration/blending with traditional practices remained intact, their means of implementation have been rapidly evolving due to the ever-expanding educational needs: enhanced vocabulary use [1, 3], new tools for interactive speech generation [3], and finally, but no less important, for speaking and pronunciation assessment [4, 5, 6, 8, 9, 10, 11, 13, 16, 17].

As has been widely covered in the recent studies of AI involvement in foreign language acquisition, the present phase of the above mentioned streams of human-computer collaboration was preceded by at least three other preliminary phases of real-life language simulation, and hence,

the unavoidable learners immersion by means of less sophisticated multimedia technologies [4, 8], use of Large Language Models [21] and a tentative use of AI chatbots for writing and interactive training.

As a matter of fact, all of the three preliminary phases outlined above have become part of the research areas for the Scientific-Methodological Laboratory of Multimedia and Digital Technologies initiated by Business Foreign Languages and Translation Department of Kharkiv Polytechnic in Collaboration with School of Foreign Languages of Vasyl Karazin National University of Kharkiv, Ukraine, for already a decade. The results and the introduction of new technologies into the educational domain have been duly presented in the prior publications [20, 21].

As one can easily trace from the above papers, the findings have always gone hand-in-hand with the blending technologies of traditional methods which are viewed and highlighted by nearly all the scholars and AI developers as indispensable. Indeed, no progress in using modern advancement in learning languages would be imaginable without educational preparation of using new technology by accompanying tutors of any kind, for that matter, despite the AI's highly personalized nature of learning. The only case of completely individualized use of AI-driven educational technologies could be the case of a learner's foreign language proficiency as high as necessary to move ahead without human tutoring. Interestingly, the same situation is true for possible completely personalized learning at the learner's required proficiency level using appropriate computer programs or multimedia technologies.

The current phase of AI penetration into the comparatively narrow field of foreign language acquisition is predominantly subtle and deep insight into the issues of simulating human voice, speech recognition and speech generation combined with their corresponding assessment. Another sidetrack would be intercultural communication and interpersonal communication in the ecolinguistic perspective. These issues have been the target of further study of AI involvement at the already paved way in doing research into the two-facet phenomenon of blended learning: through modern technologies and by means of tutor guidance.

All of the above endeavors unavoidably call for making an outline of the ever-present and constantly changing human-computer interactions, that is another major objective of the Laboratory research underway.

One more issue worth mentioning as indispensable in up-to-date efforts to harmonize technological advances and the inevitable changes in their accompanying teaching techniques is a prospective trend of involving the theories of communicative competence and ecolinguistics [21] which are bound to complete the search for human-like speech generation and assessment.

2. Related Works

The recent studies of AI-based tools adopted in language learning education are predominantly centered around human language simulation, one way or another: writing [1,2,3,6,7], speech generation [3, 6, 11, 17], speech recognition and assessment [3,5,10,13,16], and even more so for their functions as tutors and individual trainers [1, 3, 4, 5, 18, 19].

In fact, all of the above are based on AI abilities to mimic human communication in its major areas, (nonverbal is under question because of its complexity), giving way to simulating the roles of both learners and language instructors.

Even over the past few years, there has been a noticeable bias towards speaking/speech recognition and pronunciation diagnosis investigations on a par with their analysis of corresponding chatbot operations and tutoring systems [7, 8, 10, 17].

The nature of simulation and foreign language environment immersion are closely intertwined in their personalized role in training [5, 6] which is different from traditional classroom learning, while some sources claim that AI tools are even more immersive and exciting than conventional techniques [12, p. 337].

Some studies go even farther in investigating immersion techniques history, revealing at least three phases of its development: through multimedia, Large Language Models and the most up-to-date interactive tutoring and assessment [5]. It goes without saying, our present research is one more example of introducing immersion techniques in foreign language acquisition, that seems to be indispensable at least in Eastern Europe and Ukraine in particular, due to the lack of foreign language environment.

It is also stressed that collaboration between AI developers and human instructors can yield the best results [10, p. 727], for this is the human mind that initiated technological advances of AI issues in question, and this is the traditional teaching in foreign language education that may challenge AI technologies in order to foster a more beneficial and accurate instruction by AI programs that become their skilled partners, which is the case with the present findings by the Laboratory team following the experiment described below.

Overall, there's been no controversial issues in the researchers' summaries of the current and prospective collaboration of the two outlined parties, both AI developers and academics alike, as to the inevitability of their mutually beneficial advances towards creating and using the most progressive platforms for language learning education environment, despite temporary limitations and challenges of using AI tools, such as "...no enhancing students' skills in writing..."[3, p. 179], fear of teachers being replaced [4, p.13], maintaining learner motivation [6, p.208], reliability and accuracy issues [8, p. 2], pre-existing biases [9, p. 1031], "...reducing the human touch..." [15], insufficient information and teacher preparation [16].

The vast majority of the authors emphasized, though, that the needs of paramount importance require combined efforts of AI developers and educators to enhance the prospects of speedy learning. It also reveals the innovative trend to see the future of AI involvement from the perspective of the newly discovered demands from the academic community [2, 3, 4, 5, 6, 7].

In the current phase of developing AI programs for foreign language education, speech recognition [6] and language assessment are at the forefront of developers' interests [10, 13, 16, 17]. Even though AI technologies "...have been slowly embraced, ... now the attention is being focused on pronunciation improvement." [4, p. 16]. And this conclusion is in line with the experiment described in the present paper.

And last but not least, major ideas highlight the benefits and merits brought about alongside the integration of AI with conventional educational practices: smart tutoring, personalization, autonomy, lack of fear, meeting individual needs and preferences, individual pace, objectivity and time reduction [2, 3, 4, 6, 10, 11, 12, 14, 15].

The conclusive acknowledgement of AI-based education in foreign language acquisition is that of AI tools being indispensable, capable of sophisticated language guidance [5], providing increased motivation and promising results [4], enhanced proficiency [3], greater facilitation [1], and excitement [6]. However, the most persuasive observation is the fact that "...such technologies would complement the traditional learning interactions but not replace them.." [7, p. 188], which, again, is totally in line with the ideas claimed in Discussion hereby.

Another key aspect of the paper is integrating ecological considerations into human-computer interaction (HCI) which is essential to promote sustainable technology development and reduce environmental impact. As digital technologies become more pervasive, their environmental footprint (energy consumption, e-waste and resource depletion) has grown significantly. Addressing these concerns within HCI can lead to more environment friendly practices and designs.

Rethinking human-AI interactions from an ecological perspective can lead to more equitable outcomes. Research [24] suggests that adopting an ecological perspective in AI design promotes a more harmonious relationship between humans and the environment.

Integrating HCI with community citizen science initiatives can empower users to contribute to sustainability efforts. A number of studies [e.g. 25] explores how HCI can facilitate community-driven environmental research. Another highlight is designing interactions that minimize overall informational damage and encourage ecological user behavior. The studies of the "AI era" [e.g.

Fidel] discuss the evolution of sustainable interaction design, dwelling on the shift towards methodologies that consider environmental impact.

3. Methods and Materials

3.1. Research methodology

The study is based on the **hypothesis** of effective use of AI tools not only in teaching but also in the assessment of a foreign language. In the first place, the study is built on human-computer interaction (HCI) theories (3.1.1) involving the ecological aspect of this currently forming environment. Secondly it focuses on the components of the speaking skills fixed by the CEFR, which allowed conducting a detailed **experiment**. The research methodology made it possible to build an **extended model** of a human-AI ecological interaction and a **model of speaking skills** assessed by a human-AI tandem.

3.1.1. Human-AI Interaction in the Eco-Paradigm

Human-computer interaction (HCI) is concerned with understanding how people interact with technology. Primarily the focus has evolved from simply comparing humans to machines to now highlighting the dynamic relationship between the two [22]. However, the new approaches to HCI take into account ecological parameters typical for human society assessment. The new view is caused by a "technology-centered" approach that causes many failures [25]. So Qiuyu Lu et. al. [23] dwell on the ecological factor of sustainability to encompass a broader range of Sustainable Development Goals set by the United Nations. This focus helps to refine positioning within HCI, technical approaches, design strategies, evaluation methods and long-term impact. This point of view is supported by Chunchen Xu et. al. [24] who emphasize the difference between an anthropocentric and an ecological approach to HCI by advocating alternative human-AI interactions and guiding AI developments toward fostering a more caring human-ecology relationship. The newest research by Raya Fidel bridges the study of human information interaction and the design of information systems: cognitive work analysis which offers an ecological approach to design, analyzing the forces in the environment that shape human interaction with information [26].
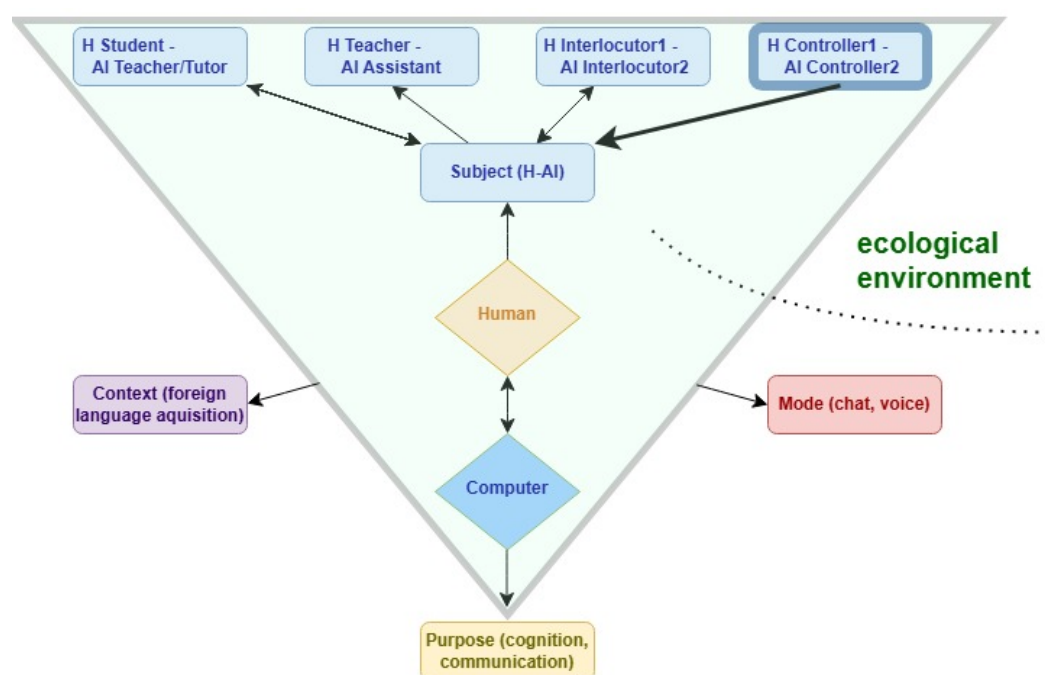
Figure 1: Model of the Human-Computer interaction with an AI role extension.

Each of the mentioned researches explores how ecological principles can enhance Human-Computer and Human-AI interactions by focusing on environmental context, sustainability, and user experience. They call for integrating ecological thinking into design to create more responsible, context-aware technologies.

Based on the new functions of the AI, we have developed an extended model of interaction between humans and AI in the framework of foreign language teaching, taking into account the departure from the anthropocentric paradigm and the focus on the ecology of the current interaction between humans and computers.

3.1.2. Speaking Competence Outline

The most common definition of "speaking" is to articulate words verbally, to communicate by means of discourse, to make a request, and to deliver a speech (Webster's New World Dictionary). The Common European Framework of Reference for Languages (CEFR) organizes speaking proficiency across six levels (A1 to C2), with specific descriptors for each level in areas such as fluency, accuracy, interaction, and range of vocabulary. This is how the CEFR addresses speech competence [27, p. 62ff]:

A1 (Beginner) level language users can produce simple phrases and sentences to express basic needs, but their competence is very limited. They are able to form basic sentences and pronounce words clearly, but accuracy and fluency may be lacking. **A2 (Elementary)** level language users can use a series of phrases and sentences to express simple opinions and needs. They can ask and answer simple questions in areas of immediate need or on familiar topics. The individual can participate in simple conversations and express basic ideas with a certain degree of fluency. **B1 (Intermediate)** can deal with most situations likely to arise whilst travelling in an area where the language is spoken; can produce simple connected text on topics that are familiar or of personal interest. Speech competence includes the ability to manage routine communication in familiar contexts, make requests, and provide explanations. **B2 (Upper Intermediate)** can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. The individual has improved fluency and can discuss a wider range of topics with greater ease and precision, making more complex statements and contributing to discussions. At the **C1 (advanced)** level, individuals can produce clear, well-structured, detailed text on complex subjects related to their field of interest, and can express themselves fluently and spontaneously without much obvious searching for expressions. At this stage, speech competence involves effective communication in a wide range of professional, academic, and social situations. At the **C2 (Proficient)** level, individuals can produce clear, smooth, well-structured discourse in complex situations, expressing themselves spontaneously, very fluently, and precisely, differentiating finer shades of meaning even in more complex situations. This marks the pinnacle of speech competence, where the individual can speak with the fluency and sophistication of a native speaker.

Some scholars restrict speaking competencies coverage to two areas, fluency and accuracy [28]. Thus, fluency deals with the student's ability to use mechanical skills, such as pauses, speed, and rhythm; speaking of accuracy, learners should pay enough attention to the exactness and the completeness of language form when speaking, such as focusing on grammatical structures, vocabulary, and pronunciation. But almost all the papers published in the "AI era" and featuring speaking competence development agree that this kind of competence involves more than just linguistic accuracy (grammar, vocabulary). They emphasize the ability to use language effectively in real-life communication, either in different contexts (social, academic, or professional) or through interaction with others.

While all speaking competence definitions highlight linguistic skills (vocabulary, grammar), some emphasize communicative competence [29, 30, 31] – the ability to interact meaningfully in various social and cultural contexts – more than others. For example, T. S. A. Sabri [29] and Mu-Hsuan Chou [32] stress the role of cultural and context-specific appropriateness in speech

competence. Additionally, Mu-Hsuan Chou introduces metacognitive awareness, suggesting that speech competence involves not only linguistic ability but also the ability to self-monitor and adjust one's speaking tasks effectively. G.S. Valdivieso-Arcos [33] and S. Sudarmo [30] place a strong emphasis on the interactional nature of speech competence, focusing on students' ability to engage in meaningful conversation and manage discourse, not just produce correct language.

All the researchers put forward the idea that speech competence in EFL learning is multifaceted, integrating linguistic proficiency, interaction skills, and the ability to navigate diverse communication contexts. However, the depth of focus varies, with some emphasizing linguistic knowledge, others highlighting interaction or cultural appropriateness, and some incorporating metacognitive strategies for self-regulation. The broadest view sees speech competence as involving a complex set of abilities that work together to enable effective communication in a foreign language. The authors' emphasis on interaction as a fundamental component of speaking skills substantiates the relevance of identifying five key aspects, which are integral to speech competence as outlined in the CEFR [27,p. 129ff]. These aspects include:

1. **Fluency**: The capacity to speak fluently, without unnecessary pauses or hesitation.
2. **Accuracy**: The correct application of grammar, vocabulary, and pronunciation.
3. **Vocabulary Range**: The diversity of vocabulary and structures employed during communication.
4. **Interaction**: The ability to engage in conversations, initiate topics, and respond appropriately.
5. **Pronunciation**: The clarity and intelligibility of spoken words.

In conclusion, the CEFR conceptualizes speech competence as the effective use of spoken language across a range of communicative contexts. It encompasses fluency, accuracy, and interactive capabilities. As proficiency levels increase, individuals' ability to participate in more complex and spontaneous communication expands, ultimately allowing them to communicate effortlessly in nearly any situation.

The hypothesis of the experiment carried out within the scope of Scientific and Methodological Laboratory at Business Foreign Languages and Translation Department in NTU "KhPI" is association with the School of Foreign Languages of the Vasyl Karazin National University Kharkiv assumes the possibility of involving artificial intelligence not only in the development of all five skills, but also in their evaluation and correction. This model was used as a research tool for conducting the experiment. It specifies the role of AI for each specific skill and was verified in the course of the experiment described below.
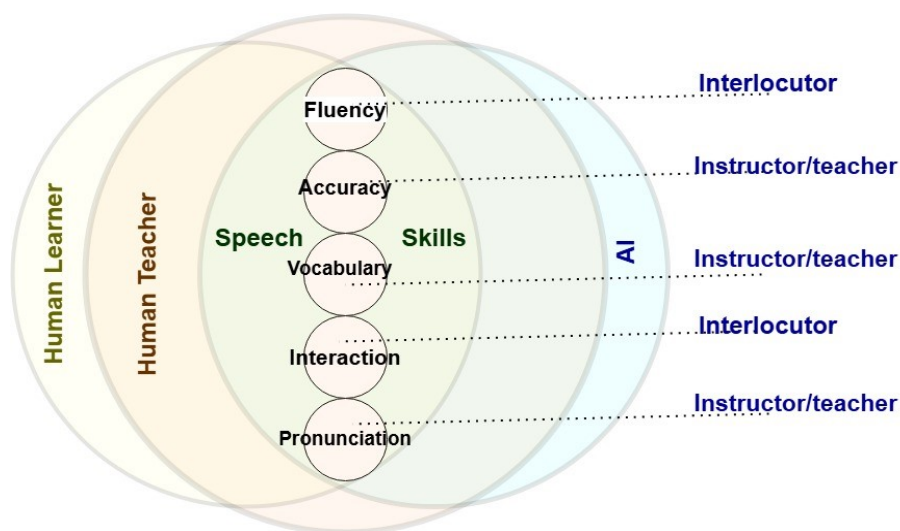
Figure 2: Model of the AI involvement into forming the speaking skills.

3.2. Research Instruments

The research instruments used in this study include:

1. AI-Powered Platform (Smalltalk2me) – This software was the primary tool for assessing students' speaking skills. It provided feedback and ratings based on five criteria: pronunciation, grammar, fluency, vocabulary, and interaction.
2. Recorded Responses – Students' spoken answers were recorded and analyzed for comparison between AI and human assessments.
3. AI-Generated Assessment Reports – These reports provided qualitative feedback, highlighting strengths and areas for improvement in students' speech.
4. Human Evaluators' Assessments – 10 human language experts also evaluated students' speaking skills, providing ratings to compare with AI-generated assessments.
5. Comparative Analysis Methods – Statistical calculations, including agreement percentage, partial agreement, overestimation and underestimation rates, were used to compare AI and human evaluations.

The constructional steps in this study rely on input data gathered through ESL students' recorded answers on the AI-powered platform Smalltalk2me during March-April of 2024. Subsequently, the second part of the experiment was carried out in November-December, 2024 when 10 human language experts were asked to estimate ESL students' speaking skills based on the recordings which had been previously assessed by AI. This part of the experiment was carried out asynchronously with the first one via Google Forms.

The authors of the research conducted both stages of the experiment, and the utilization of the collected data was done with the participants' permission.

Finally, comparative analysis methods and statistical calculations enabled the authors to explore AI vs. human assessment of ESL students' speaking skills.

3.3. The material and sample of study

The sample of this study consists of 15 second- and third-year students, aged 18-19 years old, majoring in Translation at the Business Foreign Languages and Translation Department at NTU "KhPI".

These students, whose initial level of English is on average B1-B2, voluntarily participated in the experiment which was carried out within the scope of Scientific and Methodological Laboratory at Business Foreign Languages and Translation Department in NTU "KhPI" in association with the School of Foreign Languages of the Vasyl Karazin National University Kharkiv. ESL students' English-speaking skills were assessed by the AI-powered platform Smalltalk2me and compared to human evaluations conducted by 10 language experts, aged 25+ years old, whose level of English is proficient (C1-C2).

4. The Experiment Description

As artificial intelligence continues to evolve, its application in education has sparked debates about its effectiveness compared to traditional human-led assessment. As a result, an initial experiment in the row of future prospective experiments in the area of speaking competence was conducted. It is focused on the efficiency of the role of AI as a facilitator in assessing and improving students' speaking skills and was carried out within the scope of Scientific and Methodological Laboratory at Business Foreign Languages and Translation Department in NTU "KhPI" in association with the School of Foreign Languages of the Vasyl Karazin National University Kharkiv. The obtained results prompted us to conduct further comparative analysis regarding the accuracy and effectiveness of assessment performed by a human teacher versus AI-driven software.

The experiment was performed on the AI-powered platform smalltalk2me, which is designed to be a simulator to self-practice the IELTS speaking test, job interview and everyday conversational English. The AI-driven software Smalltalk2me offers ESL students a diagnostic test to assess their speaking level and provides the feedback on areas of language acquisition improvement. Fifteen

second- and third-year students majoring in Translation at Business Foreign Languages and Translation Department in NTU "KhPI" agreed to participate in the experiment where their English-speaking skills are assessed by artificial intelligence. The recorded respondents' answers and their AI-based feedback became the focus of this research. The obtained data were collected and subsequently analyzed.

Smalltalk2me AI-driven software estimates the English speaking level based on five criteria: pronunciation, grammar, fluency, vocabulary, and interaction (See Figure 3).

SmallTalk English Speaking Level Test

Date: **18.04.2024** Duration: **9 min** Words: **1102**

Scores

Advanced ⓘ

A person at this level can understand a wide range of demanding, longer texts, and recognize implicit meaning. They can express themselves fluently and spontaneously without much obvious searching for expressions.
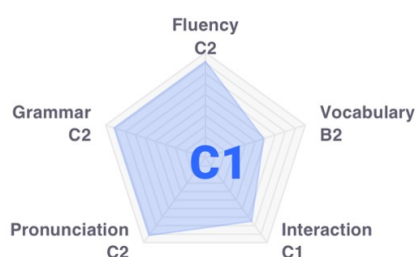
The score given may be inaccurate: skipped task #13

⌾ Certificate   ↥ Share

Fluency
C2

Vocabulary
B2

Grammar
C2

C1

Pronunciation
C2

Interaction
C1

Figure 3: A visual representation of five criteria which the Smalltalk2me software bases its speaking assessment on.

A student can pause a recording and allocate some time to think before answering the question. The format of speaking tasks is the same for all the students and the content can vary. Below, we provide examples of possible questions and tasks.
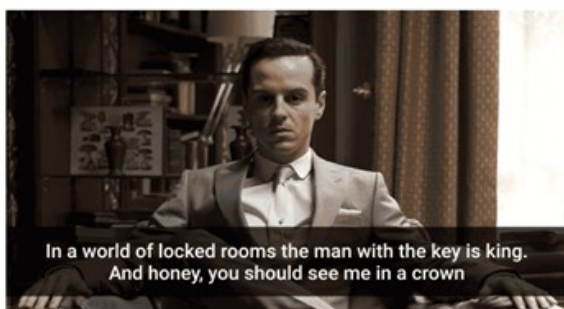
1. How are you today?
2. Where do you live? What languages do you speak?
3. Pronounce the quote from the given picture e.g. I am the one who knocks.
4. Pronounce the quote from the given picture. e.g. It wasn't logic, it was love.
5. Pronounce the quote from the given picture. e.g. In a world of locked rooms the man with the key is the king. And honey, you should see me in a crown;

6. Pronounce the quote from the given picture. e.g. I feel the same way about being a bridesmaid as you feel about Botox. Painful and unnecessary;
7. Pronounce the quote from the given picture. e.g. I'm not a psychopath, Anderson. I'm a high-functioning sociopath. Do your research. e.g. You don't trust people because they are trustworthy/ You do it because you have nothing else to rely on.
8. Read the text (about 1500 printed characters).
9. Speak for 2-3 minutes. e.g. What are your top-5 favorite websites/apps? How much time do you usually spend on the Internet? e.g. What do you usually do, when you have free time? Do you have a lot of free time during the week?
10. Here are photos from your photo album, choose one photo to describe to your friend: Speak for 2-3 minutes, in your talk remember to speak about: -Where and when the photo was taken; -What/who is in the photo; - What is happening; - Why you keep the photo in your album; -What is so special about this photo? -What emotions does it bring to you?
11. Listen to the audio and answer the questions.
12. Describe a journey that didn't go as planned. You should say: -where you were going; -who you were with; -what went wrong; -and explain what you would have done differently.
13. You were invited to your colleague's birthday party, ask your colleague questions to find out more details about their birthday party. You should ask about: - preferable type of present; -time the party starts; -the number of guests; -location.

As we can see, these tasks combine speaking, listening, and reading activities, which is important for building well-rounded communication skills. The first task is a basic conversational prompt to initiate a dialogue. It focuses on speaking and conversation initiation. It is simple and useful for warming up the student and estimating everyday language. The second task assesses the student's ability to answer personal questions and introduce themselves, conduct personal information exchange. Tasks 3-7 check pronunciation, fluency and intonation using quotes from popular media. Such tasks, like pronouncing famous quotes or describing personal photos, incorporate engaging elements, which can increase motivation by connecting language learning to students' interests and experiences. The use of popular media quotes from TV shows or movies (See Figure 4) is an effective way to engage students, especially younger ones, making learning more relevant and enjoyable.
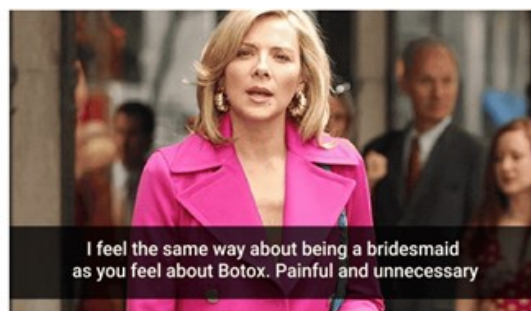
Figure 4: A visual representation of tasks with pronouncing quotes based on TV shows or movies.

Task 8 is a comprehension and fluency task, aiming to practice comprehension skills. Several tasks (e.g. speaking for 2-3 minutes, describing a photo) are designed to assess mostly speaking fluency and a range of vocabulary, which is essential for communicative competence. Listening task № 11 assesses students' understanding spoken language, which is a key aspect of communication. The task "Describe a journey that didn't go as planned" is a storytelling task that assesses a learner's ability to recount past experiences and express problems, providing opportunities for a wide range of vocabulary usage. Task 13 is designed to estimate question

formation and information gathering in a conversation, making it very applicable for real-life interactions. Overall, task complexity is variable: the first questions may be more suited for beginners, while tasks requiring longer descriptions or narratives are better for intermediate to advanced learners.

After a speaking test has been completed, AI generates assessment results based on the students' speech record instantly. The assessment report by AI begins with such complementary phrases as: "Confidence. Native would understand 95% of your speech", "Jargon King. Phrasal Verbs are your strong points", "Accuracy Achiever. It's amazing how many of your sentences are correct!", "Grammar Expert. 52 % of your sentences have complex structure or advanced constructions", "Synonyms Master. We are impressed by your synonyms variety", "Coherence Genius. The use of linking words is excellent!", "Vocabulary Nailer. We love your active vocabulary", "Story Teller. Story telling is your skill. You came up with an amazing answer!" and so on. As we can see, in such way AI establishes good rapport with students quickly and draws their attention to their strengths in terms of coherence of their speech, grammatical accuracy, self-confidence, narrative skills and a plethora of vocabulary. To begin with, Smalltalk2me platform gives a positive feedback on students' speaking and highlights areas to improve (See Figure 5). Absence or scarcity of grammatical errors, presence of advanced grammar constructions such as conditionals or relative clauses, using phrasal verbs and exploiting a high-level active vocabulary in the respondent's speech are recognized by AI as areas of excellence. AI algorithms count the percentage of complex structures in sentences, generate possible synonyms and offer linking words to point to possible aspects of speaking which should be improved. Thus, offering encouragement is an essential constituent part of feedback quality and its impact on student learning. Consequently, a complementary section is intrinsically imbedded in AI learning platforms.

Nicely done

✅ You made very few **grammatical errors.**

✅ You used 8 advanced **grammar constructions**: reported speech, relative clauses, passive voice, going to, tenses, conditional, modals, phrasal verbs.

✅ You're an expert at using **phrasal verbs**: "be take in", "find out", "get up", "go on", "relate to".

✅ You have a large **active vocabulary** and use more **high-level words** than most speakers of English do.

Things to improve

⚠️ Only 35% of your sentences had a **complex structure**.

⚠️ Use some **synonyms**: "photo", "take", "people", "place", "lot" were used too often when talking about the same thing.

⚠️ Add various **linking words**, such as 'besides', 'to start with', 'in spite of', 'to my mind', 'in case of', 'therefore'.

Figure 5: AI feedback on strengths and areas of improvement following student's speaking test conducted on the AI-powered platform smalltalk2me.

In the next section of the assessment, AI identifies sentences that do not sound naturally and offers a more appropriate variant under the heading "How a native speaker would say the same", for example the excessively formal sentence "Receiving bad news daily about conflicts in my country is distressing" is offered to be replaced with a more informal one "Hearing about daily

conflicts in my country is saddening”. The used vocabulary is divided according to CEFR levels in percentage terms and the examples are cited (See Figure 6).

Vocabulary statistics
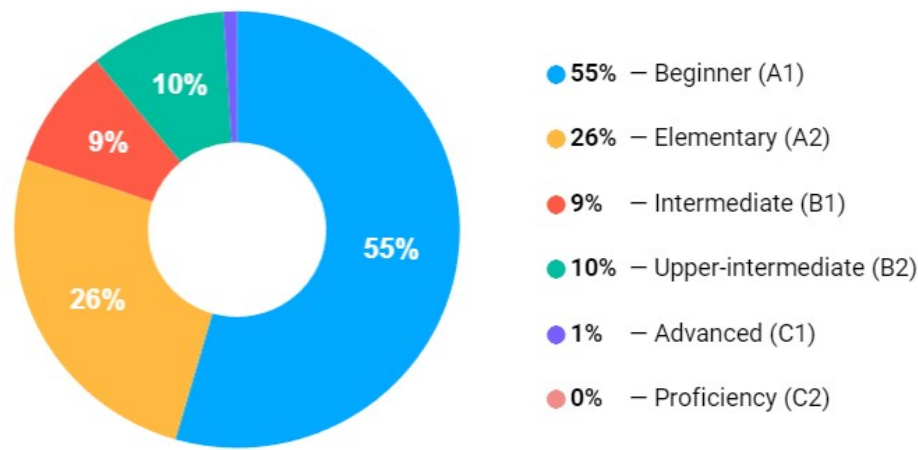
Figure 6: The example of analyzing the student's vocabulary after passing a diagnostic test in the application "SmallTalk2Me".

In terms of pronunciation and fluency, AI ranks the speed of speech (words per minute): 90-150 – common speaking rate for native speakers in the US, less than 90 – too boring, more than 150 – too fast to understand (See Figure 7).

Speaking rate (words per minute)
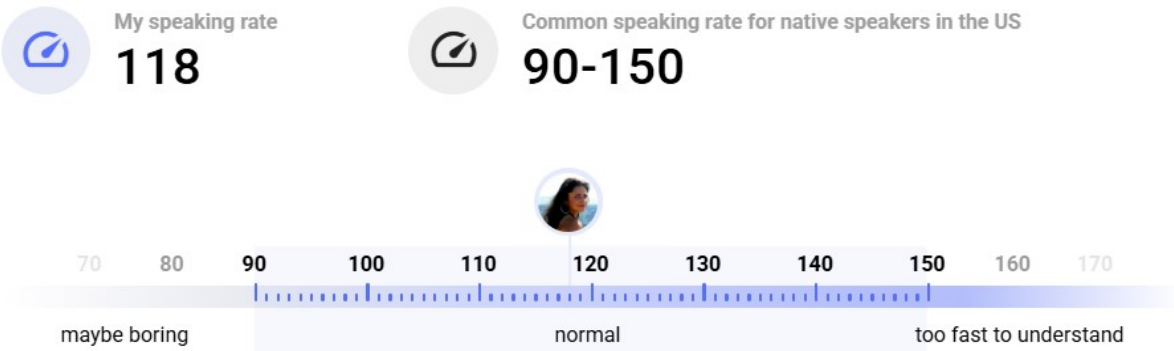
Figure 7: Defining the student's speaking rate by AI-driven platform smalltalk2me.

Along with identifying the speaking rate, AI detects the student's pronunciation mistakes and phonetic inaccuracies that change the meaning and gives the report on them with the correct pronunciation to practise.

Regarding grammar, AI divides the mistakes made into categories, for example "Article", "Verb constructions", "Prepositions", calculates their number in the speech and gives references to the rules in an online Cambridge Dictionary.

Hence, we can sum up that AI-driven software relies on predefined algorithms and large datasets to evaluate ESL student speaking performance. This ensures consistency and eliminates personal biases, gives instant feedback with extensive explanations and comprehensive progress statistics.

The following stage of the given experiment was to conduct assessment of the same students' recordings by human language experts using the same criteria. Human language experts listened to the students' records and filled in the Google forms with the following questions:

1. Which grammar mistakes have you noticed in the student's analyzed speech?
2. Which pronunciation mistakes have you noticed in the student's analyzed speech?
3. How would you assess the level of student's speech according to each criterion: Grammar (A2-C2), Fluency (A2-C2), Vocabulary(A2-C2), Interaction(A2-C2), Pronunciation(A2-C2)?
4. How would you assess the overall level of student's speaking skills (A2-C2)?
5. Which strengths would you point out in the student's speech:

   - "Confidence. Native would understand 95% of your speech",
   - "Jargon King. Phrasal Verbs are your strong points",
   - "Accuracy Achiever. It's amazing how many of your sentences are correct!",
   - "Grammar Expert. 52 % of your sentences have complex structure or advanced constructions",
   - "Synonyms Master. We are impressed by your synonyms variety",
   - "Coherence Genius. The use of linking words is excellent!",
   - "Vocabulary Nailer. We love your active vocabulary",
   - "Story Teller. Story telling is your skill. You came up with an amazing answer!",
   - "Your variant"?

6. What aspects of speaking can be marked in the student's speech as "Nicely done"?
7. What aspects of speaking can be marked in the student's speech as "Things to improve"?
8. Have you noticed any repetitions of words? If yes, which ones?
9. What synonyms would you offer to the student instead of too simple or repetitive words which have been used?
10. How would you mainly characterize the level of overall vocabulary which has been used by the student? (A2-C2)
11. The student's speaking rate and pausing are:

    - below native speaker's level,
    - normal,
    - too fast to understand.

So, each student has been assessed by five criteria using level descriptors (from A2 to C2). Additionally, the dataset includes qualitative feedback, identifying strengths, areas for improvement, and synonyms for repetitive/simple words. The dataset (See Figure 8) contains students' speaking assessments with ratings for Grammar, Fluency, Vocabulary, Interaction, Pronunciation, and Overall Level. Each cell contains two levels separated by a hyphen (e.g., "C1-C2"), where the first value represents the human assessment and the second represents the AI assessment.

5. Results

Calculating agreement percentages and checking for significant differences helped us analyze how well AI matches human ratings according to Figure 8 below.

The agreement between human and AI assessments here refers to how often AI gives the same rating as a human for a particular skill. In this case, agreement was calculated as the percentage of students for whom AI assigned the exact same level as the human evaluator. For example, in Grammar, AI and humans gave the same rating for 25% of students, meaning that in 75% of cases,

AI's rating was different from the human's. Calculations show that the agreement between human and AI assessments varies across categories:

- Interaction (37.5%) has the highest agreement.
- Grammar, Fluency, and Vocabulary (25%) show moderate agreement.
- Pronunciation (18.75%) and Overall Level (12.5%) have the lowest agreement.

This suggests that AI mostly struggles with assessing **overall speaking level** and **pronunciation**, while it performs relatively better in evaluating **interaction skills.**

	A	B	C Grammar	D Fluency	E Vocabulary	F Interaction	G Pronunciation	H Overall Level
1	№		Grammar	Fluency	Vocabulary	Interaction	Pronunciation	Overall Level
2	1	Student 1	C1-C2	B1-C1	B1-B2	B2-C1	B2-C1	B2-C1
3	2	Student 2	B1-C1	C1-C1	B1-B2	B2-C1	B1-C1	B1-B2
4	3	Student 3	C2-C1	B1-B2	B2-B2	C1-C1	C1-C2	B2-C1
5	4	Student 4	C1-C2	C1-C2	B2-C1	C1-C2	C1-C2	C1-C1+
6	5	Student 5	B2-C1	B1-B2	B1-B2	B1-C1	B2-C2	B1-B2+
7	6	Student 6	B2-C1	B1-C1	B2-B2	B1-C1	B2-C2	B1-C1
8	7	Student 7	B1-C2	B1-C1	B2-C2	B2-C2	B2-C2	B1-C1+
9	8	Student 8	A2-B2	B1-B1	A2-B1	B1-B2	A2-C1	A2-B1+
10	9	Student 9	B1-B1	B2-C1	B1-B1	B2-B2	B1-C1	B1-B2
11	10	Student 10	B1-C1	B1-B2	A2-B1	B1-B2	B2-B2	B1-B2
12	11	Student 11	C1-C2	C2-C2	C2-B2	C1-C1	C1-C2	C2-C1
13	12	Student 12	C2-C2	C2-C1	C1-B2	C1-C1	C2-C2	C2-C1
14	13	Student 13	B2-C1	C1-B1	B2-B1	B2-B2	B1-C1	B2-B2
15	14	Student 14	A2-C1	A2-B2	A2-B2	A2-B2	A2-B1	A2-B2
16	15	Student 15	B2-B2	B2-C1	B1-B2	B2-C1	B2-C2	B2-B2+
17			human vs AI	human vs AI	human vs AI	human vs AI	human vs AI	human vs AI

Figure 8: Summary Data from Human vs. AI Assessment of Students' Speaking Skills

These findings lead to calculating **partial agreement**, which includes cases where AI's rating is either exactly the same as the human's or one level higher/lower**.** This will give a broader picture of how closely AI's assessment aligns with human judgment. Here are the partial agreement percentages (cases where AI was either exactly the same as the human rating or only one level different):

- **Vocabulary (81.25%)** and **Interaction (75%)** have the highest alignment.
- **Grammar (68.75%)** and **Fluency (68.75%)** also show strong agreement.
- **Pronunciation (50%)** and **Overall Level (56.25%)** have the lowest agreement, suggesting AI struggles mostly in these areas.

This indicates that AI often comes close to human ratings, even if it doesn't match exactly**.**

Here is the visualization (See Figure 9) comparing exact agreement (AI matches human ratings exactly) and partial agreement (AI is within one level of human ratings). The blue bars represent exact matches, while the red bars show how often AI's assessment is close but not identical.

This highlights that AI performs best in Vocabulary and Interaction, while Pronunciation and Overall Level remain the most challenging.
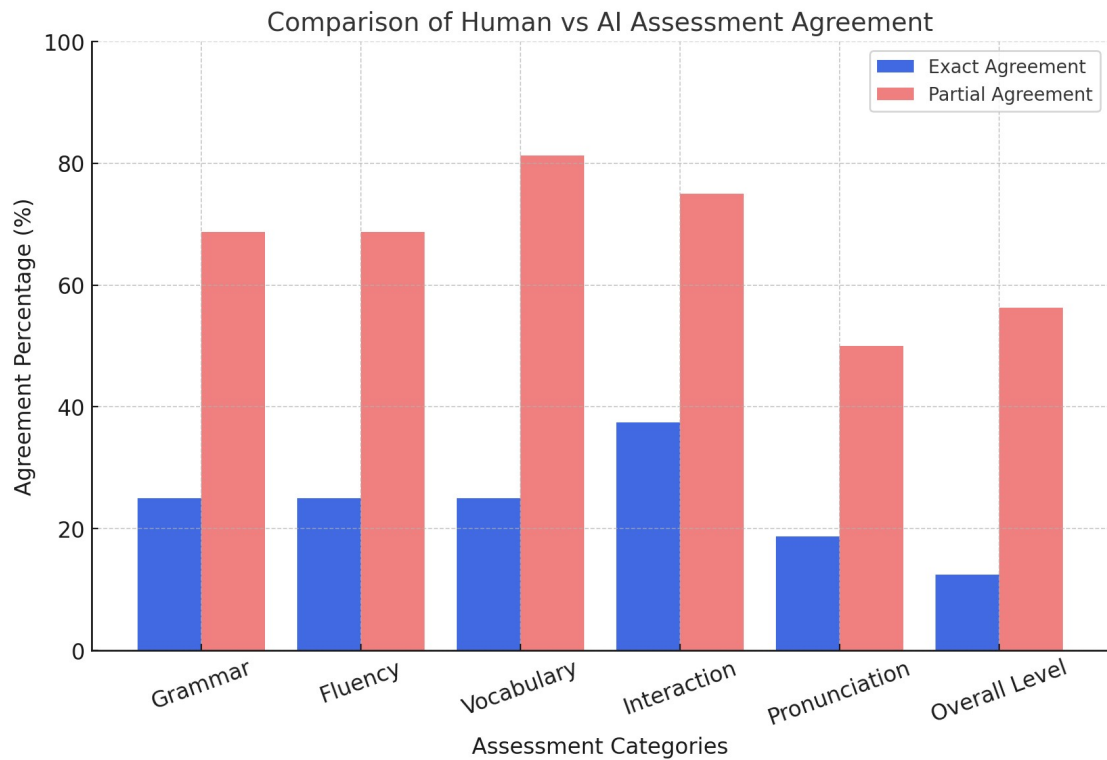
Figure 9: Comparison of Human vs AI Assessment Agreement

Further analysis is focused on overestimation and underestimation, checking how often AI rates students higher or lower than human evaluators.

Here is how AI tends to overestimate or underestimate student levels compared to human evaluators:

✔ AI consistently overestimates:
- Pronunciation (81.25%) is the most overestimated skill.
- Grammar (68.75%), Fluency (62.5%), and Interaction (62.5%) also have high overestimation rates.
- Vocabulary (56.25%) and Overall Level (43.75%) show moderate overestimation.

✔ AI rarely underestimates:
- Interaction and Pronunciation (0%) are never underestimated.
- Grammar (6.25%), Fluency (12.5%), and Overall Level (12.5%) have minor underestimation.
- Vocabulary (18.75%) is the most underestimated skill, suggesting AI sometimes rates students lower than human evaluators.

A visual representation of these data is shown in Figure 10. Here is the visualization comparing AI overestimation (orange) and AI underestimation (green) across different assessment categories. We can see that AI strongly overestimates Pronunciation (81.25%) and Grammar (68.75%), rarely underestimates students, except for Vocabulary (18.75%) and Overall Level (12.5%) and never underestimates Interaction and Pronunciation, meaning it consistently rates them equal to or higher than human evaluators.
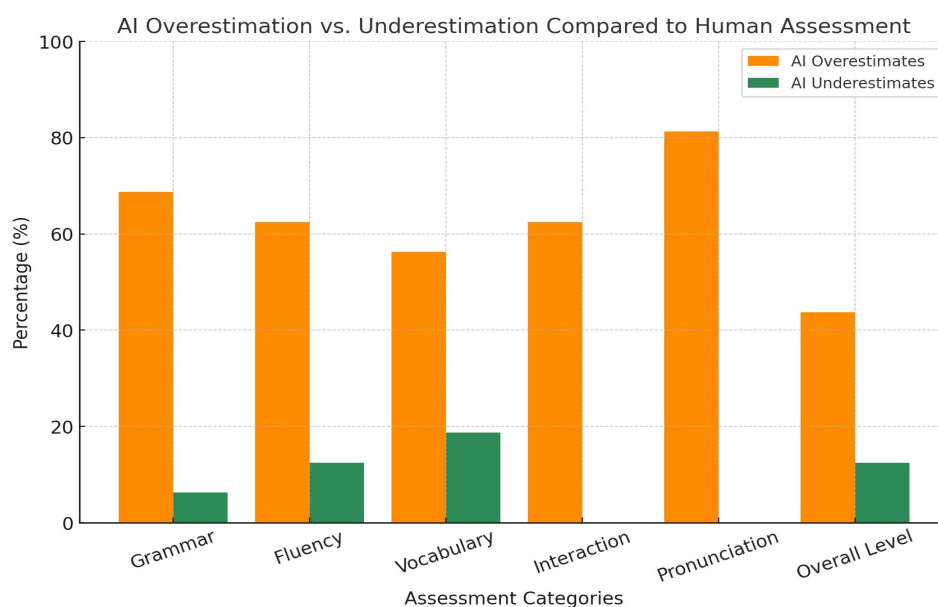
Figure 10: AI Overestimation vs Underestimation in relation to human evaluators' assessment

Our further insights are focused on a breakdown of which specific levels AI tends to overrate or underrate (See Figure 11). We analyzed which proficiency levels (e.g., A2, B1, B2 etc.) AI overestimates or underestimates most frequently. This shows if AI is biased toward rating certain levels higher or lower than humans.
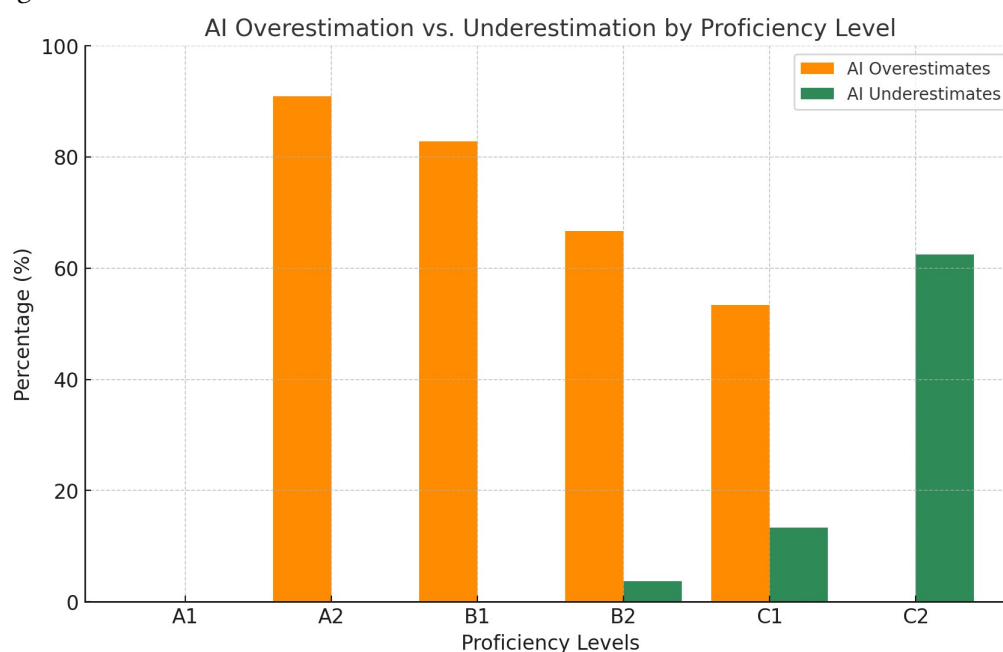
Figure 11: AI Overestimation vs. Underestimation by Proficiency Level

Here is how AI tends to overestimate or underestimate different proficiency levels:

AI heavily overestimates A2 (90.9%) and B1 (82.8%), meaning students at these levels were often rated higher by AI than by humans.

B2 (66.7%) and C1 (53.3%) are also frequently overrated, but with slightly more balance.

AI severely underestimates C2 (62.5%), meaning students at this level were often rated lower than they should be.

6. Discussion

As is evident from the purpose of the study and the outcoming results, that the objective of the experiment presented hereby is the comparison of AI-based speaking assessment results to those by human experts, while the core of the comparison being qualitative approach of accuracy and objectivity. The data provided claim the overall unanimity of assessment by the opposing counterparts as to their ratings of the students' proficiency from B2 to C2, however, not exceeding one grade within a separate speaking category, where, e.g. human C1 is paired with AI C2, but no further. We term them as partial agreements.

Nonetheless, the assessment grades within categories Vocabulary, Grammar, Pronunciation, Interaction, Fluency and Overall Level vary significantly. The highest alignments of the two counterparts rest with Vocabulary and Interaction. Less alignment, although still strong, can be traced within Grammar and Fluency, whereas the lowest one lies within Pronunciation and Overall Level, to suggest that these categories are either more difficult to grasp and process by AI tools, or the obvious discrepancy is caused by other human factors (the level differences rising from 1 to 2 in the discussed areas.) The discrepancies might be also associated with vulnerability of Pronunciation and Overall level (which, again, includes Pronunciation) as the most intangible and subjective for assessment categories.

As far as the general Comparison of Human vs AI Assessment Agreement goes, the discrepancies between exact and partial agreements within the categories confirm the above observation of Pronunciation and Overall Level being the most vulnerable and subjective areas due to their intangible blurred nature, sometimes hard to grasp either by the presently developed AI tools for this kind of assessment or due to the multifaceted nature of human assessment, including quite a number of factors like human experts' proficiency level, biased to different variants of spoken English (British or American), experience in the field and overall objectivity.

Another revelation worth mentioning is difference in overestimation and underestimation by AI tools in terms of students' proficiency levels: A1, B1, B2 are more likely to be overestimated, while underestimated higher levels (C1 and C2) can be partially explained by AI putting higher demands to higher proficiency levels.

It's worth stressing the fact that the previous developments in AI advancement for foreign language acquisition were predominantly concentrated on writing, grammar and reading skills rather than speaking and especially pronunciation assessment. The latter two, though, are an integral part among the other five key aspects of speech competence (Fluency, Accuracy, Vocabulary Range, Interaction and Pronunciation), as fundamental components of speaking skills. In this study we term them intangible in contrast to more tangible writing, grammar and reading.

Alongside the above results, the experiment once more confirmed the incessant need to integrate the efforts of AI developers and human instructors in order to achieve the required results in all the areas mentioned above. With time, the focus of these efforts will inevitably turn to the still barren land of other sophisticated subtle intangible domains of psycholinguistics and AI identification of social roles, interpersonal communication and already mentioned rules of cross-cultural communication, as well as the most recently tackled ecolinguistics.

As illustrated in the diagrams presented in Figure 12, media competence in 2019 slightly exceeded 2 points. This modest level is consistent with the general principle that any skill, regardless of its nature, requires time to develop. A significant increase to 8.48 points was observed in 2020, with the Covid-19 pandemic playing a pivotal role. The pandemic, along with the global shift to online environments due to quarantine restrictions, accelerated this growth.

In 2021, a further increase in media competence was noted, though it was more gradual. This can be attributed to well-established educational principles: as individuals solidify their foundational knowledge in any field, they transition from novices to professionals. Subsequently, their continued professional development may result in less dramatic improvements, unless there is a fundamental shift in the subject matter or the profession itself.

Despite the challenges presented by the onset of war in 2022, the growth of media competence persisted, albeit at a slower pace. During this period, the competence rose by nearly half a point, underscoring the resilience of the educational environment. The shift to a fully online format, particularly in the frontline region of Kharkiv, Ukraine, under continuous danger to its residents did not diminish this upward trajectory.

As shown in Figure 12, the media competence saw a sharp increase between 2019 and 2020, primarily due to the introduction of new digital platforms. However, the growth slowed in the subsequent three years, with only modest increases, and even a slight decline in 2023. This reduction is likely to be attributed to the overwhelming number of new platforms, which may surpass individuals' capacity to process information efficiently. Nevertheless, survey results from participants in the experiment in 2024 indicate a slight increase in their confidence when interacting with various forms of AI. This can be explained by the accumulation of experience over time, which enhances individuals' familiarity with and ability to navigate these technologies.

Self estimated level of media competence
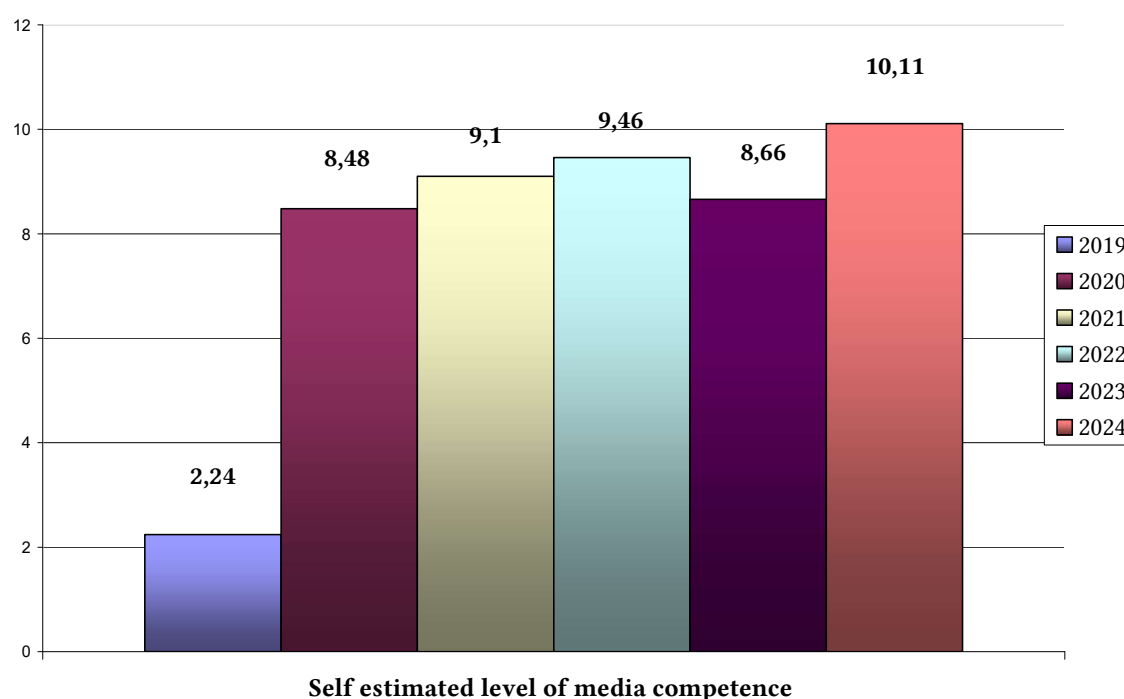
Figure 12: Dynamics of testees' media competence in 2019-2024

7. Conclusions

The current investigation into the realm of present-day digital technologies attempts to outline the perspectives of AI coverage of all the areas of real-life communication, their penetration into the already integrated socio- and psycholinguistics, among them communicative competence, cross-cultural communication and ecolinguistics. As a natural and quite expected phenomenon, technology is gradually but continuously blended with conventional teaching methods in foreign language acquisition, and yields undeniable beneficial results with AI leading the way in learners immersion and real-life language simulation with their predecessors in the form of multimedia technologies and Large Language Models.

Nonetheless, in the question "Who is the teacher?", the advances in the area under question are not necessarily due to the rapid technological breakthrough, but sometimes due to the blended approach, where educational needs foster the appearance of unprecedented discoveries and pave the way for innovative AI tools, never heard of before. One of such tentative pushes for the AI

program developers' efforts, hopefully, might be the results of the experiment presented in this publication.

It is obvious, though, that the present phase of AI development, aimed at speaking assessment, has to concentrate around the challenges for AI tools in their consistency, objectivity and accuracy based on the new areas of research: variability of English across borders, the prosodics of pronunciation (tone and emotional coloring included), as well as some cultural features and ecolinguistics.

The field of HCI has undergone significant evolution, transitioning from a simplistic comparison between humans and machines to a more sophisticated understanding of their dynamic interplay. Recent research highlights the growing importance of ecological parameters, sustainability, and environmental context in shaping both HCI and human-AI interactions. The extended model of human-AI interaction for foreign language learning presented in this paper serves as a prime example of this ecological perspective, moving beyond traditional anthropocentric paradigms to emphasize the broader ecology of human-computer interaction. By integrating ecological thinking, this model seeks to foster the development of technologies that are not only context-aware but also promote sustainability and responsible use. Ultimately, the incorporation of ecological principles into HCI marks a transformative shift towards designing technologies that are not only effective but also socially and environmentally responsible.

The logic of the sequence of the authors' previous studies determined the object of this paper - the assessment of productive language competence of speaking by human experts and an AI platform. Speaking competence is a multifaceted skill that encompasses various dimensions, including fluency, accuracy, vocabulary range, interaction, and pronunciation, as outlined in the CEFR framework. While traditional definitions have often emphasized linguistic accuracy, contemporary research increasingly underscores the significance of communicative competence, cultural appropriateness, and metacognitive strategies for effective communication. The integration of AI in language learning, as demonstrated in the experiment conducted by the Scientific and Methodological Laboratory at NTU "KhPI," illustrates AI's potential not only to develop but also to assess and rectify these critical components of speaking competence. This innovative approach underscores the evolving role of technology in enhancing language education, opening new avenues for cultivating effective communication skills across diverse contexts.

Declaration on Generative AI

In the course of this paper preparation , the authors used smalltalk2me in order to evaluate the speaking skills of the testees and GPT-4o to generate bar charts "Comparison of Human vs AI Assessment Agreement", "AI Overestimation vs Underestimation in relation to human evaluators' assessment" and "AI Overestimation vs. Underestimation by Proficiency Level" (Figures 9-11) based on the obtained data in the course of the experiment. DeepL Write was used to improve writing style. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] A. Kholili, A. Afandi, AI-Powered Writing Tools: Does Word-tune Bring Benefits for EFL Learners' Writing Performance?, MARAS: Jurnal Penelitian Multidisplin, Vol. 2, No. 3, September 2024, 1345-1352.

[2] Hansol Lee, Jang Ho Lee, The effects of AI-guided individualized language learning: A meta-analysis, Language Learning & Technology, Volume 28, Issue 2, June 2024, 134–162

[3] Yuanyuan Xiong, A Review of the Features and Efficacy of Chat-GPT AI Writing Assistant in Influencing EFL Learners' English Writing Skills, Transactions on Social Science, Education and Humanities Research, 11, 177-183. https://doi.org/10.62051/bxz3th29

[4] H. Vančová, AI and AI-powered tools for pronunciation training, Journal of Language and Cultural Education, 2023, 11(3), 12–24. doi: 10.2478/jolace-2023-0022

[5] A. S. E. AbuSahyon, A. Alzyoud, O. Alshorman, and B. Al-Absi, AI-driven Technology and Chatbots as Tools for Enhancing English Language Learning in the Context of Second Language Acquisition: A Review Study, IJMST, Vol. 10, No. 1, Oct. 2023, 1209-1223. https://doi.org/10.15379/ijmst.v10i1.2829

[6] D. Kristiawan, Bashar Y., K., and D. A. Pradana, Artificial intelligence in English language learning: A systematic review of ai tools, applications, and pedagogical outcomes, The Art of Teaching English as a Foreign Language (TATEFL), 5(2), 2024, 207-218. https://doi.org/10.36663/tatefl.v5i2.912

[7] L. Gutiérrez, Artificial Intelligence in Language Education: Navigating the Potential and Challenges of Chatbots and NLP, Research Studies in English Language Teaching and Learning, 1(3), 2023, 180–191. https://doi.org/10.62583/rseltl.v1i3.44

[8] M. Zhu, and Chaoran Wang, A Systematic Review of Artificial Intelligence in Language Education from 2013 to 2023: Current Status and Future Implications. Preprint, Posted: 4 Jan 2024. http://dx.doi.org/10.2139/ssrn.4684304. URL: https://ssrn.com/abstract=4684304

[9] S. Devi, A. S. Boruah , S. Nirban , D. Nimavat, K. K. Bajaj, Ethical Considerations in Using Artificial Intelligence to Improve Teaching and Learning, Tuijin Jishu/Journal of Propulsion Technology, Vol. 44, No. 4, 2023, 1031–1038. https://doi.org/10.52783/tjjpt.v44.i4.966

[10] H. U. A. I. Al-Abbas, H. H.Halim, and N. N. Nurjati, Harnessing the use of artificial intelligence in language assessment: A systematic comprehensive review. Tell-Us Journal, 9(3), 2023, 723–745. doi:10.22202/tus.2023.v9i3.7366

[11] D. Abimanto, and W. Sumarsono, Improving English Pronunciation with AI Speech-Recognition Technology. Acitya: Journal of Teaching and Education, 6(1), 2024), 146–156. https://doi.org/10.30650/ajte.v6i1.3810

[12] J. A. Mananay, Integrating Artificial Intelligence (AI) in Language Teaching: Effectiveness, Challenges, and Strategies, International Journal of Learning, Teaching and Educational Research, Vol. 23, No. 9, September 2024, 361–382. https://doi.org/10.26803/ijlter.23.9.19

[13] A. K. Betal, Enhancing Second Language Acquisition through Artificial Intelligence (AI): Current Insights and Future Directions, Journal for Research Scholars and Professionals of English Language Teaching, Vol. 7, Issue 39, September 2023. https://doi.org/10.54850/jrspelt.7.39.003

[14] M. S. Fountoulakis, Evaluating the Impact of AI Tools on Language Proficiency and Intercultural Communication in Second Language Education, International Journal of Second and Foreign Language Education, 3 (1), 2024. 12–26. https://doi.org/10.33422/ijsfle.v3i1.768

[15] M. J. K. O. Jian, Personalized learning through AI, Advances in Engineering Innovation,5, 2023.16–19. https://doi.org/10.54254/2977-3903/5/2023039

[16] Jin Ha Woo, and Heeyoul Choi, Systematic Review for AI-based Language Learning Tools, Preprint submitted to JDCS, October 2021. doi:10.48550/arXiv.2111.04455

[17] A. E. Cahyono, and R. Rosita, The impact of using Ai-based language learning platforms on English speaking skills of college students, Transtool, Volume 2, No. 2, 2023, 1–8. https://doi.org/10.55047/transtool.v2i2.1352

[18] R. Rusmiyanto, N. Huriati, N. Fitriani, N. Tyas, A. Rofi'i, and M. Sari, The Role Of Artificial Intelligence (AI) In Developing English Language Learner's Communication Skills. Journal on Education, 6(1), 2023, 750–757. https://doi.org/10.31004/joe.v6i1.2990

[19] A. A. Al Harbi, The Uses of Machine Learning (ML) in Teaching and Learning English Language: A Methodical Review, Journal of Education, January-Part 3-(93), 2022, 26–52. doi:10.21608/edusohag.2022.212355

[20] O. Cherednichenko, O. Yanholenko, A.Badan, N. Onishchenko, N. Akopiants, Large Language Models for Foreign Language Acquisition, Proceedings of the 8th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2024), Volume IV: Computational Linguistics Workshop, Vol-3722, 2024, 101-130. URL: https://ceur-ws.org/Vol-3722/

[21] A. Badan, N. Onishchenko, and O. Zeniakin, Digital Technologies for Communication Simulation in Foreign Language Learning under Pandemic, Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Systems, volume 3171 of COLINS-2022, Gliwice, Poland, 2022, 1160–1180.

[22] S. Schleidgen, O. Friedrich, S. Gerlek, et al., The concept of "interaction" in debates on human–machine interaction, Humanit Soc Sci Commun, 10, 551 (2023). https://doi.org/10.1057/s41599-023-02060-8

[23] Qiuyu Lu, A. Danielescu, V. Iyer, P. Lopes, Lining Yao, Ecological HCI: Reflection and Future, CHI EA '24: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 2024, Article No.: 577, 1–4. URL: https://doi.org/10.1145/3613905.3643985

[24] Chunchen Xu, Xiao Ge, AI as a Child of Mother Earth: Regrounding Human-AI Interaction in Ecological Thinking, Conference: CHI '24: CHI Conference on Human Factors in Computing Systems, 2024, Article No.: 546, 1–9. URL: https://doi.org/10.1145/3613905.3644065

[25] Wei Xu, M. Dainoff, Enabling Human-Centered AI: A New Junction and Shared Journey Between AI and HCI Communities, Interactions, 2023, Volume 30, Issue 1, 42–47. https://doi.org/10.1145/3571883

[26] R. Fidel, Human Information Interaction. An Ecological Approach to Information Behavior, MIT Press, 364 p.

[27] Council of Europe, Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume, Council of Europe Publishing, Strasbourg, 2020. URL: https://www.coe.int/lang-cefr

[28] T. Colle, What is a speaking skill? Definition and competencies, Tenrycolle.com, September 29, 2022. URL: https://tenrycolle.com/skills/speaking/what-is-a-speaking-skill-definition-and-competencies/

[29] T. S. A. Sabri, Communicative competence in English as a foreign language. International Journal of English Language Teaching, Vol.6, No.9, 2018, 27-33. URL: https://eajournals.org/wp-content/uploads/Communicative-Competence-in-English-as-a-Foreign-Language.pdf

[30] S. Sudarmo, The importance of speaking in English as a foreign language between skillful and thoughtful competencies: studying sociolinguistics perspectives, Linguistics and Culture Review, 5(S1), 2021, 113-124. https://doi.org/10.37028/lingcure.v5nS1.1321. URL: https://media.neliti.com/media/publications/500746-the-importance-of-speaking-in-english-as-972308c0.pdf

[31] D. N. Tallibayev, Methods used in speech development of English language learners, American journal of language, literacy and learning in STEM education, Volume 01, Issue 04, 2023. URL: https://grnjournal.online/index.php/STEM/article/download/1636/1637/3311

[32] Mu-Hsuan Chou, An Integrated Approach to Developing and Assessing EFL Students' Speaking Competence, Language Education & Assessment, 4 (1), 2021, 19–37. https://doi.org/10.29140/lea.v4n1.428 *ERIC*. URL: https://files.eric.ed.gov/fulltext/EJ1295081.pdf

[33] G.S. Valdivieso-Arcos, J.C.Argudo-Serrano, Improving EFL students' speaking skill through academic oral expression tasks at the high school level, Episteme Koinonia. Revista Electrónica de Ciencias de la Educación, Humanidades, Artes y Bellas Artes, Año V, Vol. V, N°9, Enero – Junio, 2022. http://dx.doi.org/10.35381/e.k.v5i9.1663. URL: https://ve.scielo.org/pdf/ek/v5n9/2665-0282-ek-5-09-56.pdf