

Aleksandar Kavgić*
Filozofski fakultet
Univerzitet u Novom Sadu

UDK 371.3::811.111]:004.738.5
811.111'276.6:339.138
DOI: 10.19090/gff.v49i2.2534
orcid.org/0000-0001-5065-7084
Original scientific paper

EFFECTS OF USING AI IN TEACHING ENGLISH: A WORKPLACE-ENGLISH CASE STUDY

This quantitative pilot case study-into workplace English (WE) teaching practice set out to determine the extent to which the use of artificial intelligence (AI) in creating teaching materials can influence the teaching outcomes and the student experience. The study employed a traditional experiment methodology with the control and the experimental group who were given a pretest and a posttest to determine both the extent of language skills and the level of learning experience satisfaction. The experiment was conducted in a small IT company among adult learners of English, comprising individuals from three teams, totaling 37 people who were split into two groups. The control group of 18 people was instructed according to a standard textbook by OUP, Business Results, combined with International Edition of English for Customer Care, as well as with general English-in-use materials created by the instructor. The experimental group of 19 people was lectured using custom teaching materials created by AI, which was instructed to prepare tailor-made learning resources based on email exchanges, phone calls, social media posts previously created or recorded within the IT company. The AI was also instructed to create test materials (questions and quizzes) that were focused on each individual piece of the content generated in this way. The results of the experiment indicate that both the control and the experimental group have improved the language skills which were the focus of the learning materials in a statistically significant manner, whereby the control group did not show statistically significant improvements among individuals, while the experimental group showed statistically significant improvements among the majority of individuals. Furthermore, the experimental group reported higher level of satisfaction after the course end. In sum, the results of this small-scale pilot study seem to indicate that the use of AI in preparation of learning resources for teaching WE leads to better acquisition of relevant skills as well as a more positive, subjectively measured, learning experience.

Key words: AI, workplace English, teaching, learning materials, learning experience

* kavgi@ff.uns.ac.rs

1. INTRODUCTION

On November 30, 2022, OpenAI unveiled a preliminary demonstration of its AI solution, ChatGPT, which rapidly gained popularity on social media platforms as users began to share its capabilities: the wide range of examples and stories shared by early adopters encompassed extremely diverse tasks, ranging from organizing trips to crafting tales and programming software, so that in just a span of five days, the chatbot had garnered the attention of over a million users (Marr, 2023). In under 18 months, as of April 2024, ChatGPT has grown to have over 180 million users and 1.6 billion visits per month (Duarte, Fabio, 2024).

Naturally, with such an unprecedented growth in user base, the AI capabilities were employed to virtually all types of tasks involving human languages, including teaching and learning foreign languages. In truth, the idea of using AI in foreign language learning is not new and was floated in the early 2000s, whereby AI was proposed as a framework to enable teaching English as a foreign language (e.g. J. Jia, 2009). Naturally, with the wide availability of the technology, the last 18 months have seen an increase in scientific publications exploring the use of AI in language learning, ranging from general ones that are focused on professional and societal impact (e.g. Firat, 2023; Fitria, 2021; Meyer et al., 2023; Opyr et al., 2022; and Rukiati et al., 2023), to more focused ones that represent case studies on how to integrate AI into the classroom and facilitate students' interaction with it (e.g. Dewi et al., 2021; F. Jia et al., 2022; and Lee et al., 2023). Nonetheless, these studies generally address the integration of AI as an interactive participant, or an intelligent tool, into the learning/teaching process in real-time or via online learning platforms and apps, while this study sets out to explore the benefits of using AI in a more indirect way, as a means to produce high-quality teaching and learning materials that are customized for a particular target group of learners of workplace English (WE), while retaining the traditional teaching methodology, i.e. keeping AI out of the classroom in realtime, but using it as an input for teacher and student interactions.

Therefore, this study set out to investigate if the AI, which is constantly being advertised as an augmentation of human cognitive abilities that boosts productivity up to several orders of magnitude, can be used by EFL/WE teachers to prepare personalized language learning materials that have sufficient quality to achieve the educational goals and are not detrimental, but hopefully beneficial, to the learning experience by the students. The remaining part of the study is

organized so that, first, in Section 2, “Theoretical Considerations: Personalized Language Learning Materials and the Learning Experience” as well as the overall EFL/WE domain of the study are presented, which is then, in Section 3, followed by a detailed overview of the “Methodology” used to conduct the study, leading to the presentation of “Research Findings” in Section 4, and, naturally, “Concluding Remarks” in Section 5.

2. THEORETICAL CONSIDERATIONS: PERSONALIZED LANGUAGE LEARNING MATERIALS AND THE LEARNING EXPERIENCE

Given the overarching interest in assessing the potential benefits, or drawbacks, of using the AI in preparing personalized instruction materials to increase the quality of the learning experience, this section has two main parts, which function as theoretical foundations of the study: the first part provides a general overview of the importance of personalized instruction materials, while the second part introduces the notion of learning experience, in particular in relation to aforementioned personalized instruction materials. In addition to functioning as a theoretical anchor of the study, these sections also provide additional reasons and rationale why the two were chosen as focal points of the study.

2.1. The importance of personalized teaching and learning materials (PTLM)

Why were the personalized teaching and learning materials (PTLMs) chosen as the focus of the study? Generally, PTLMs in English as a Foreign Language (EFL) or WE contexts are essential for the teaching/learning process because they can significantly enhance the effectiveness of instruction by addressing the unique needs, interests, and cultural backgrounds of learners, because when teaching materials are tailored to reflect the interests as well as linguistic abilities and learning styles of individual students, they can boost motivation, facilitate better engagement, and lead to improved language retention (e.g. Nunan, 1988). Moreover, PTLMs can help incorporate relevant cultural references, ranging, in the domain of IT, from latest IT events and news to more subtle references revolving around industry-specific jargon and scenarios such as client requirement gathering or application of prevalent methodologies like Agile and SCRUM, that resonate with the learners' own experiences, making the learning process more meaningful and contextually relevant. This approach aligns with the principles of learner-centered education, which is widely recognized for its benefits in promoting significant and deep learning experiences (e.g. Fitria,

2022; and Lucas et al., 2018).

However, the development of such personalized EFL materials is a highly demanding task that requires substantial time and effort from educators, as well as careful planning in terms of estimating how long it will take to create such materials (e.g. Peachey, 2013). Crafting PTLMs involves a thorough analysis of learners' needs, preferences, and proficiency levels, which must be meticulously integrated into the design and adaptation of teaching materials: this process often entails continuous revision and customization to align with evolving educational goals and learning trajectories, but also to align with workplace/contextual interests of students, all of which is explained from multiple perspectives by a multitude of authors in e.g. Tomlinson (2013). The labor-intensive nature of this approach can be claimed to place considerable strain on teachers, especially those with large classes or limited resources, potentially leading to a compromise in the quality of teaching, as well as potential burnout or overwork issues on the part of the teacher, which is the reason why, to a large extent, teachers generally rely on ready-made teaching and learning materials in the form of coursebooks.

Despite these challenges, the investment in personalized learning materials is generally supported by research as beneficial. Studies such as those by Tomlinson and Masuhara (2013) and Nguyen (2013) emphasize the positive impact of individualized learning strategies on language acquisition, highlighting how tailored approaches cater to both specific and general learner needs, thus enhancing overall educational outcomes. It was exactly for this reason that the following research question was posed and deemed worth the effort of conducting a small-scale pilot study: can AI be used to create PTLMs, thus increasing the quality of learning for students and reducing the amount of effort by the teacher.

2.2. The concept of learning experience

Shifting to a bird's eye perspective, effective EFL teaching strategies generally incorporate a learner-centered approach, where instruction is tailored to the individual needs, interests, and cultural backgrounds of the students, thus promoting a more personalized and impactful learning experience, whereby PTLMs represent one of the means of achieving effective and learner-centered EFL experience. In other words, the use of PTLMs is inherently tied to the concept of the learning experience in EFL, which encompasses a broad range of methodological approaches to interactions and activities that are designed to facilitate the acquisition of English by non-native speakers in the most optimal and idiosyncratic way. This experience is not confined merely to the conventional

classroom setting but extends to a more immersive engagement where learners interact with the language through multiple modalities, including multimedia resources, peer interactions, and real-world practice and examples. Through it all, the focus is on creating a supportive learner-centered environment that encourages language use in the most relevant communicative contexts, thereby enhancing both fluency and comprehension through the application of various methodologies. For instance, Richards and Rodgers (2014) provide comprehensive insights into various teaching methodologies that enhance the EFL learning experience, ranging from alternative approaches, such as total physical response, multiple intelligences, (now controversial) neurolinguistic programming and competency-based language teaching, to current communicative approaches, such as content-based instruction, which can be said to be the best-suited method of teaching in the context of development and application of PTLMs.

Regardless of the method applied, the importance of the learning experience in EFL cannot be overstated. It plays a crucial role in language retention and the practical application of English in diverse contexts, which are essential for academic, professional, and personal growth because engaging learning experiences boost motivation, which is a critical factor in the long-term success of language learning. A detailed overview of psychological and social aspects of language learning, with a particular emphasis on the importance of a supportive learning environment that includes personalized learning materials, can be found, for example, in Lightbown and Spada (2021). Additionally, when learners see tangible progress and can use the language to achieve real-life goals, their engagement and enthusiasm increase, leading to deeper and more sustained learning: this has been well-known for several decades and has been addressed as early as the 1990s, for example, by Nunan (1999), who, mostly in the form of a personal account, described the struggles in applying EFL theories and concepts into one's teaching practice and exploring in detail, the practical applications for promoting active learning and learner autonomy, including aspects of what is here referred to as PTLMs.

In summary, there is a large volume of works that collectively underline the significance of thoughtful teaching practices in the EFL context that lead to a better student engagement, highlighting the effectiveness and implications of enriching the learning experience in EFL, where learner-centered and personalized approaches, such as PTLMs, have been identified as one of the means to achieve the goal of improved learning experience. It should be noted that the literature review does not engage with papers that target the integration of AI into the development/production of PTLMs or language teaching materials in general, but

instead only highlights the importance of PTLMs in general, because, at the time of the publication, there were no publications, as far as the author's search abilities go, that deal with PTLMs in the domain of AI.

3. METHODOLOGY: THE STUDY'S METHODOLOGICAL DESIGN AND AI INTERACTIONS FOR PTLM PRODUCTION

This section is organized in two parts. First, the methodology of the study is presented. Then, some examples of creating PTLMs are presented through brief description and several screenshots of interactions with AI, i.e. OpenAI ChatGPT 4. At this point, it should be noted that "prompt engineering", i.e. the process of devising instructions and provision of language inputs for ChatGPT 4 requires a study in its own right and cannot be presented here due to a different focus (and wordcount limits) of this paper, nor can a detailed overview of learning materials be provided, for the same reason. The methodology described in the following two sub-sections was applied during in-house WE course at the company where the author worked short-term. This teaching experiment took place between November 2023 and the first week of March 2024, at a pace of one 60-minute class per week (a total of 17 classes per group).

3.1. The methodological design of the study

In this small-scale pilot / case study, a quasi-experimental design was employed to assess the efficacy of AI-generated PTLMs in improving language proficiency among employees at a small IT company in Novi Sad. The company has requested to be anonymous, but has no objections to the publication of this study as long as its anonymity is maintained.

Specifically, the participants were 37 company employees from three different departments taking in-house workplace English classes. All individuals were independently tested by the company's HR department using the Oxford Quick Placement Test, where all employees were at least B2 (according to the CEFRL). The participants were divided into two groups based on their availability for scheduled class times, resulting in a natural, as well as random, allocation to either the control group or the experimental group. The control group, consisting of 18 individuals, received instruction through a standard curriculum using established textbooks from OUP listed in sources, at the end of the paper, and additional learning materials on hypothetical expressions (including subjunctive),

conditionals and conjuncts, mostly adapted from general university courses on English grammar and/or online sources (i.e. general English in use, with no special focus on the company relevant topics). Conversely, the experimental group, comprising 19 participants, engaged only with custom-tailored teaching materials generated by AI. These personalized teaching and learning materials were crafted using prompts fed into an AI system where the AI was first provided with samples of the company's in-house text material, whereby the AI then produced content closely mirroring real-life communication scenarios encountered by the learners in their daily work activities, such as email exchanges, blogs, and transcripts of teleconference calls focusing on common professional challenges like providing customer support, articulating the value proposition, discussing pricing and other work issues falling within the scope of regular work and company operations. The process of creating AI-generated PTLMs is briefly described in the next section.

Assessment of language proficiency was conducted using two pre- and two post-tests that measured: a) user experience and b) mastery of specific language-in-use topics included in the course: hypothetical expressions (including subjunctive), conditionals and conjuncts. The language-in-use tests, previously validated in university-level courses by the same teachers delivering the WE course, allowed for a comparative analysis of learning outcomes between the two instructional approaches. Notably, both groups underwent identical testing procedures to ensure consistency in evaluation. The pre-test and post-test data and scores were statistically analyzed, as described in the following chapter. The methodological approach aimed to isolate the educational impact of AI-driven content customization, providing insights into its potential advantages over traditional learning materials in a professional setting.

3.2. AI interactions used for PTLM production

In this study, AI, specifically OpenAI's ChatGPT4, was utilized to engineer highly personalized learning materials for the experimental group. The process employed a more sophisticated AI interaction mechanism based on expanding prompts that integrated user inputs from company-specific sources.

The process began with a collection of textual data from the company's own communications, including blogs, emails, teleconference transcripts, brochures, website pages, application notes, technical documentation and other publicly available, or customer accessible, documents. These inputs served as a foundation for the AI to understand and replicate the company's unique linguistic identity, including its tone, voice, and specialized jargon. By feeding this data into

the AI system, the study leveraged the advanced natural language processing technology to generate educational content that was not only tailored to the specific needs of the learners, but also resonated with the authentic communicative style of the company while dealing with relevant company topics.

The use of AI in this manner allowed for a creation of an adaptive teaching-and-learning-material creation environment where materials evolved in response to the specific requirements of participants, marking a significant shift from the one-size-fits-all approach typical of traditional textbooks. In the experimental group, AI-generated materials focused on key business English topics extracted from textual data collected from the company's communications, where the AI was instructed to modify it so as to include hypothetical expressions (including subjunctive), conditionals and conjuncts, which were, in this particular course, identified as the set of English-in-use skills to be upgraded for the participants' roles in customer support, sales, and business development.

For example, in Figure 1, one can see the initial prompt to create a fill-in-the-gap exercise for conditionals on the basis of a new company blog (not shown in Figure 1 due to spatial limitations (the blog is ~1200 words), and indicated with three dots in the screenshot), which was given to the AI as an input in a previous prompt. In Figure 2, it can be seen how the same blog is used to ask ChatGPT to prepare a reading comprehension task which was used at the beginning of the class. It is worth noting that ChatGPT can also create the key to all tasks and can adjust the level of task, e.g. make exercises correspond to C1-level language skills, which is exemplified in Figure 3 and Figure 4. Finally, based on this experiment, it can be also mentioned that it is better to keep multiple PTLMs on the same topic in the same conversation with ChatGPT, as this allows the AI engine to contextualize and interlink tasks, using the same topics and terminology that are recognized by WE students and correspond to their regular work topics.

As indicated above, this personalized content, i.e. PTLM, was the focal point of the study, as it was expected to facilitate a deeper understanding and more practical application of the language, thereby improving both the learners' language skills and their professional competencies. Specifically, this AI-driven pedagogical approach was hypothesized to not only enhance language proficiency, but also increase learner satisfaction by providing a more individualized and contextually relevant learning experience, setting the stage for a comprehensive analysis and evaluation of AI's potential in professional language training, which is the focus of the following chapter.

You
Hi ChatGPT! Can you create a fill in the gap exercise on the use of conditionals and hypothetical grammatical constructions for C1 level of English on the basis of the blog which I will provide in the next prompt? I would need a task with 10 gaps to be filled, with conditional or hypothetical expressions.

ChatGPT
Of course! Please provide the blog text or a specific excerpt from it, and I'll create a fill-in-the-gap exercise focusing on conditionals and hypothetical constructions suitable for a C1 level of English.



ChatGPT
Based on the provided blog text, here's a fill-in-the-gap exercise focusing on conditionals and hypothetical constructions for a C1 level of English:

Fill in the Gaps Exercise: Conditionals and Hypothetical Constructions

Read the text below and fill in the gaps with the appropriate conditional or hypothetical expressions.

1. If Prof. Adriano had not embraced HIL technology, he _____ (not be able to) share such detailed insights into power system protection.
2. _____ (If/Should) you visit the CESPE lab, you will see an impressive array of HIL setups.
3. Had the university not partnered with Typhoon HIL, they _____ (not achieve) such significant advancements in their research.
4. _____ (If) more universities adopted HIL simulation, they _____ (enhance) their engineering programs significantly.
5. Without the use of HIL, the researchers _____ (struggle) more with real-world application issues.
6. _____ (Were) a student to ignore the importance of this technology, he or she _____ (miss out on) valuable hands-on experience.
7. By the end of the course, students who have engaged thoroughly with HIL technology _____ (be) better prepared for industry challenges.
8. _____ (If) Prof. Adriano _____ (not introduce) digital twins in his teachings, students might not understand the complexities of real-world scenarios.
9. Should there be any technical issues during the simulations, the existing setup _____ (allow) for immediate troubleshooting.
10. _____ (If) we had not conducted an interview with Prof. Adriano, we _____ (not realize) the full scope of UFSM's contributions to grid modernization.

Figure 1: A basic prompt for AI (ChatGPT 4) for generating a PTLM focused on language use: occasionally, the AI response requires correction, e.g. in 6) where the answer is already given in the brackets (were): all responses were manually checked and corrected when necessary.

**You**

Thank you ChatGPT! Can you also create a general reading comprehension task with five questions about the blog? Each question should be a multiple-choice question with 5 options, where one option should be a clear distractor.

**ChatGPT**

Certainly! Here's a reading comprehension task with multiple-choice questions based on the provided blog text.

Reading Comprehension Task

Read the blog text about Prof. Adriano and HIL simulation at UFSM, then answer the following questions. Choose the best answer from the options provided.

1. What role does Prof. Adriano Peres de Morais hold at the Federal University of Santa Maria (UFSM)?
 - A) Assistant Professor
 - B) Guest Lecturer
 - C) Full Professor
 - D) Visiting Scholar
 - E) Research Assistant
2. What is one of the main benefits of using HIL simulation in power system protection research?
 - A) It simplifies administrative tasks.
 - B) It reduces risk and costs in real-world applications.
 - C) It replaces the need for digital substations.
 - D) It eliminates the use of artificial intelligence.
 - E) It focuses solely on theoretical aspects.
3. According to the blog, what does the CESPE lab at UFSM integrate into its HIL environment?
 - A) Only Typhoon HIL devices
 - B) Only GPS systems
 - C) Various relays and Intelligent Electronic Devices (IEDs)
 - D) Mainly theoretical models
 - E) Exclusively software simulations
4. How has the partnership between UFSM and Typhoon HIL evolved?
 - A) It has remained purely theoretical without practical engagement.
 - B) It began with joint projects and remained limited to local engagements.
 - C) It started from conference collaborations and developed into joint projects.
 - D) It is restricted to occasional consultations.
 - E) It only involves student exchanges.
5. What educational outcome does Prof. Adriano emphasize from using HIL in his courses?
 - A) Decreasing the relevance of traditional learning methods.
 - B) Providing students with hands-on experience using modern technology.
 - C) Shifting entirely to online courses.
 - D) Focusing solely on power system analysis.
 - E) Eliminating the need for physical classrooms.

Figure 2: A follow-up prompt for AI (ChatGPT 4) for generating a PTLM focused on reading comprehension.

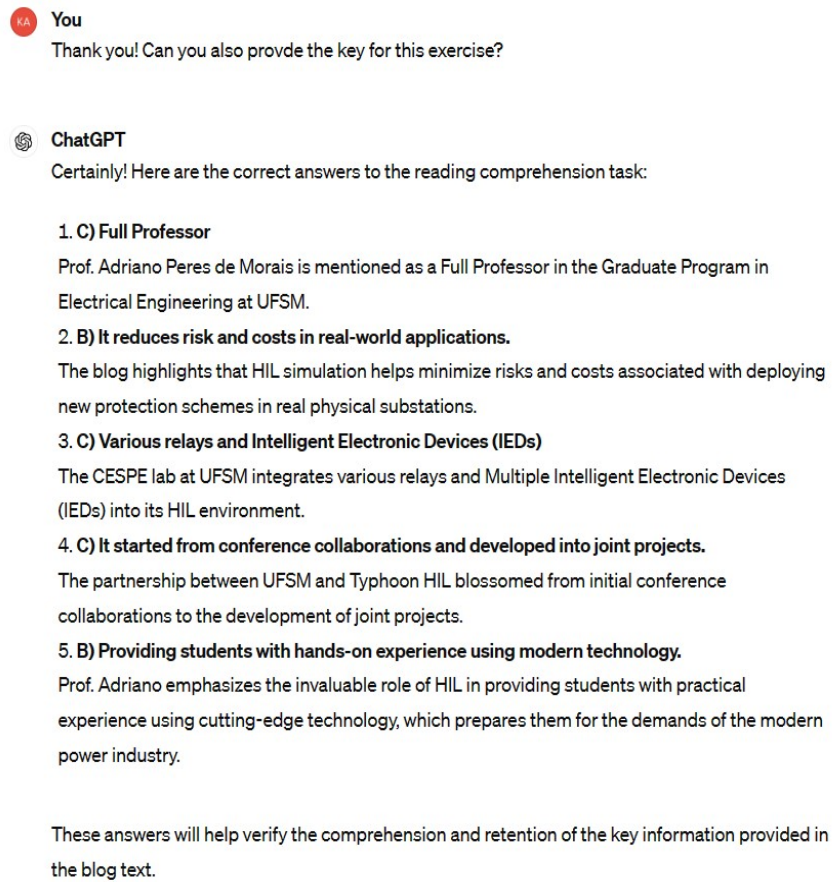


Figure 3: An additional prompt for AI (ChatGPT 4), which generates the key for the PTLM focused on reading comprehension.

 You

Much appreciated! Could you also provide the key for the first exercise, i.e. fill in the gap exercise?

 ChatGPT

Certainly! Here are the correct answers for the fill-in-the-gap exercise on conditionals and hypothetical constructions:

1. would not have been able to

If Prof. Adriano had not embraced HIL technology, he **would not have been able to** share such detailed insights into power system protection.

2. Should

Should you visit the CESPE lab, you will see an impressive array of HIL setups.

3. would not have achieved

Had the university not partnered with Typhoon HIL, they **would not have achieved** such significant advancements in their research.

4. If, would enhance

If more universities adopted HIL simulation, they **would enhance** their engineering programs significantly.

5. would struggle

Without the use of HIL, the researchers **would struggle** more with real-world application issues.

6. Were, would miss out on

Were a student to ignore the importance of this technology, he or she **would miss out on** valuable hands-on experience.

7. would be

By the end of the course, students who have engaged thoroughly with HIL technology **would be** better prepared for industry challenges.

8. If, had not introduced

If Prof. Adriano **had not introduced** digital twins in his teachings, students might not understand the complexities of real-world scenarios.

9. would allow

Should there be any technical issues during the simulations, the existing setup **would allow** for immediate troubleshooting.

10. If, would not realize

If we had not conducted an interview with Prof. Adriano, we **would not realize** the full scope of UFSM's contributions to grid modernization.

Figure 4: A prompt for AI (ChatGPT 4) which shows the importance of keeping all prompts for a single PTLM in the same chat.

In Figure 4, one should also notice that the key for fill-in-the-gap PTLM can be generated at any point in the human-AI interaction, as the AI is “aware” of

For individual students, virtually the same formula was used as for the groups, but it was applied for each column, e.g. for the student whose scores were inserted in column B, the formula was =T.TEST(B2:B4, B5:B7, 2, 1), i.e. the T.TEST formula operated on a single column, instead of sets of columns. The results for the control group and for the experimental group are presented in Table 2 and Table 3, respectively. As it can be seen from the bottom row of Table 2, no student in the control group increased their use of language skills in a statistically significant way, but, as shown in Table 3, 14 out of 19 students had statistically significantly better scores in the posttest. This can be taken as a strong indication that the use of AI generated PTLMs had a significant impact on acquisition of language skills, as opposed to individual results in the control group.

CONTROL GROUP																		
QUESTION:	CG01	CG02	CG03	CG04	CG05	CG06	CG07	CG08	CG09	CG10	CG11	CG12	CG13	CG14	CG15	CG16	CG17	CG18
PRE TEST Hypothetical expressions:	5	4	5	6	5	3	6	5	8	4	6	7	5	4	6	5	4	6
PRE TEST Conditionals:	6	6	5	4	5	5	7	5	7	5	7	8	6	5	6	6	6	8
PRE TEST Conjunctions:	5	5	4	5	5	4	7	7	10	5	6	8	5	7	6	6	7	7
POST TEST Hypothetical expressions:	6	4	5	6	5	4	5	5	7	5	7	7	5	5	5	4	3	6
POST TEST Conditionals:	7	7	7	5	6	5	8	5	8	5	7	9	7	5	8	6	7	9
POST TEST Conjunctions:	5	6	5	5	5	5	7	8	10	7	7	9	6	6	5	7	8	9
The difference between each pair of pre-test and post-test scores	-0.67	-0.67	-1.00	-0.33	-0.33	-0.67	0.00	-0.33	0.00	-1.00	-0.67	-0.67	-0.67	0.00	0.00	0.00	-0.33	-1.00
the standard deviation of the differences	0.75	1.11	0.90	0.69	0.37	0.75	0.94	1.21	1.25	0.90	0.47	0.82	0.75	0.94	1.00	0.94	1.77	1.26
T-TEST results (T-TEST EXCEL FUNCTION)	0.18	0.18	0.23	0.42	0.42	0.18	1.00	0.42	1.00	0.23	0.18	0.18	0.18	1.00	1.00	1.00	0.67	0.23
VERDICT:	NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.NOT.SIG.																	

Table 2: Use-of-English score for the pre-test and post-test in the control group for each individual (NOT.SIG = not a statistically significant difference)

EXPERIMENTAL GROUP																			
QUESTION:	EXG01	EXG02	EXG03	EXG04	EXG05	EXG06	EXG07	EXG08	EXG09	EXG10	EXG11	EXG12	EXG13	EXG14	EXG15	EXG16	EXG17	EXG18	
PRE TEST Hypothetical expressions:	5	5	5	6	6	7	6	4	5	5	5	6	7	6	6	5	7	6	7
PRE TEST Conditionals:	5	6	5	7	7	8	7	5	6	5	5	4	8	8	7	5	8	7	7
PRE TEST Conjunctions:	6	8	7	8	6	6	9	5	5	7	5	6	8	7	7	7	8	6	8
POST TEST Hypothetical expressions:	7	9	8	9	9	9	9	8	8	8	7	8	8	7	9	7	9	8	10
POST TEST Conditionals:	8	10	9	10	9	10	10	8	9	8	8	7	8	10	9	8	8	8	9
POST TEST Conjunctions:	8	9	10	10	10	10	9	10	9	10	9	9	9	7	10	9	10	8	10
The difference between each pair of pre-test and post-test scores	-2.33	-3.00	-3.33	-2.67	-3.00	-2.33	-2.33	-3.67	-3.67	-2.67	-3.00	-2.67	-0.67	-1.00	-2.67	-2.33	-1.33	-1.67	-2.33
the standard deviation of the differences	1.26	1.77	1.89	1.49	1.57	1.34	1.50	1.89	1.95	1.53	1.61	1.60	0.58	1.26	1.41	1.46	0.94	0.90	1.26
T-TEST results (T-TEST EXCEL FUNCTION)	0.02	0.10	0.01	0.02	0.04	0.02	0.07	0.01	0.03	0.02	0.04	0.02	0.18	0.23	0.02	0.02	0.18	0.04	0.02
VERDICT:	SIGNIFIC.NOT.SIG.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.NOT.SIG.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.SIGNIFIC.																		

Table 3: Use-of-English score for the pre-test and post-test in the experimental group for each individual (NOT.SIG = not a statistically significant difference, SIGNIFIC. = a statistically significant difference)

To confirm this positive effect of AI-generated PTLMs in the experimental group, the Welch's t-test was used for all three language skills in the control and the experimental group, but the formula was adjusted to not assume equal variances. This approach makes it possible to evaluate the differences between the control and experimental groups across each skill without the assumption that the data from both groups have similar spreads. As can be seen in Table 4, the differences between the control and experimental group are statistically significant, allowing us to reject the null hypothesis that PTLMs have no effect on better acquisition of language skills. On the contrary, Table 4 shows that AI-generated PTLMs as the only learning intervention in the experimental group do seem to provide statistically better acquisition of language skills.

THE TEST	RESULTS OF THE TEST
TTEST Hypothetical expressions:	1.53328E-10
TTEST Conditionals:	1.27871E-05
TEST Conjuncts:	2.92094E-06

Table 4: *Use-of-English Welch's T-test score (not assuming equal variance) for the pre-test and post-test for each language feature covered by the course: comparing the control group and experimental group test scores.*

4.2. The effects of using AI-generated PTLMs on learning experience

The learning experience responses were analyzed using the TextBlob sentiment analysis tool (Beevi et al., 2024). TextBlob sentiment score is a polarity measure which ranges from -1 (most negative) to +1 (most positive). A score of 0 generally indicates a neutral sentiment, where it is worth noting that some responses, such as "Yes" or "Definitely," while intended to be positive, may not carry explicit sentiment-laden words, and thus are scored as neutral (0.0) by the sentiment analysis tool (ibid.). The sentiment analysis results are presented in Table 5 below for each response individually. It should be noted that, due to spatial limitations, this analysis focuses on a single post-test learning experience question – "Did this course meet your expectations?" – while the entirety of the learning experience questionnaire (shown in Figure 5) will be further analyzed in a separate paper.

PRE-TEST EXPECTATIONS AND EXPERIENCE:
Group:
Age:
Gender:
Occupation:
Educational background:
What are your expectations from this course?
What specific skills or knowledge do you hope to gain?
How relevant do you think the course content is to your job or everyday work? (1=irrelevant, 5=extremely relevant)
Have you attended similar courses before? If yes, how satisfied were you with those experiences?
How do you prefer to learn (e.g., lectures, hands-on activities, group discussions)?
POST-TEST EXPECTATIONS AND EXPERIENCE:
Did this course meet your expectations?
What specific skills or knowledge did you gain?
How relevant do you think the course content is to your job or everyday work? (1=irrelevant, 5=extremely relevant)
Have you attended similar courses before? If yes, was this one better or worse than the good previous one(s)?

Figure 5: Pre- and post-test learning experience questions.

When it comes to the control group, the average sentiment score is 0.117, indicating a mildly positive sentiment overall. The standard deviation is 0.195, which shows some variation in the strength of the positive sentiment. Having said that, the minimum sentiment score is -0.1, which is slightly negative, but this may be an outlier since the 25th percentile is at 0, indicating that most responses are neutral to positive. Furthermore, the 50th percentile (median) is at 0, which means half of the responses are neutral in sentiment, while the maximum sentiment score is 0.5, showing a more strongly positive sentiment in few responses. On the whole, these scores suggest that the majority of the control group had a neutral, but mostly positive experience with the course, with it meeting or somewhat exceeding their expectations in most cases.

On the other hand, the average sentiment score in the experimental group is 0.298, which is positive, and a score which is 152.8% higher than in the control group. Standard deviation is 0.336, which indicates a range of sentiment strengths. In comparison to the control group, there were no negative sentiments and the minimum sentiment score is 0.0 (neutral), which is also the 25th percentile score, i.e. at least a quarter of responses are neutral. The median sentiment score is 0.1, which is a positive sentiment, while the maximum sentiment score is 0.91, which is a very strong positive sentiment. It should be mentioned that the response

"Abso-fucking-lutely" has a strong positive sentiment in colloquial terms, but the sentiment analysis tool does not recognize this as positive and assigns a neutral score (0.0). This illustrates a limitation of sentiment analysis tools when it comes to slang or expressions that deviate from standard language. Furthermore, responses such as "It was good", "I loved it", "Yes, it was very good", and "It was amazing!" suggest a very high level of satisfaction, contributing to the higher overall positive sentiment in the control group.

Overall, these results from Table 5 show that the experimental group had a more positive reception of the course when compared to the control group, with several responses indicating a very positive sentiment that significantly exceeds mere satisfaction.

CONTROL GROUP	EXPERIMENTAL GROUP
Definitely: 0.0	I was positively surprised: 0.10
Better than expected: 0.5	Yes: 0.00
Absolutely: 0.2	Better than the expectations: 0.50
Hell yeah: 0.0	It was good: 0.70
Yes: 0.0	Yes: 0.00
Exceeded them: 0.0	I loved it: 0.70
Yes and more than that: 0.25	Better than what I expected: 0.20
Yes, definitely: 0.0	Yes: 0.00
Beyond expectations: 0.0	Of course: 0.00
Yes, absolutely: 0.2	Yep: 0.00
Mostly exceeded them: 0.0	Yes, it was very good: 0.91
Yes: 0.0	Yes: 0.00
It did: 0.0	Yeah: 0.00
Yes: 0.0	Surely: 0.50
Beyond what I expected: 0.0	Mostly: 0.50
Sure thing: 0.5	So so: 0.00
I was positively surprised: 0.25	Kind of: 0.60
Yes: 0.0	It was amazing!: 0.75
	Abso-fucking-lutely!: 0.00

Table 5: *Learning experience responses from the control and experimental group as analyzed using the TextBlog sentiment analysis tool*

5. CONCLUDING REMARKS

The case study presented in this paper explored the efficacy of AI-generated, personalized teaching and learning materials (PTLMs) in enhancing workplace English (WE), i.e. a subtype of business English, proficiency among a small group of IT professionals, compared to traditional textbook-based and/or general-purpose materials. The findings indicate that the use of PTLMs tailored specifically to the learners' professional contexts and generated via AI, significantly improved language proficiency, as evidenced by the superior performance of the experimental group in both individual and group-level assessments. These results underscore the potential of AI to transform educational content creation, making it more relevant and effective for learners by aligning it closely with their real-world communicative and workplace needs.

Moreover, the study highlighted the importance of personalized learning environments in fostering higher levels of learner engagement and satisfaction. Despite the same effort by the instructor and similar levels of participation and homework submission across both study groups (which was not discussed earlier, due to spatial limitation), the experimental group reported a higher degree of satisfaction. This suggests that the use of PTLMs, i.e. the customization of learning materials to meet specific professional needs, not only enhances learning outcomes but also improves the overall learning experience, potentially leading to increased motivation and better retention rates.

However, the study's limitations, including its small scale and the singular context of a single IT company, mean that these findings should be interpreted with caution. Future research should be conducted with the goal to replicate this study across various industries and with much larger participant groups to confirm the findings and explore the broader applicability of AI in WE/Business English educational settings. Additionally, comparing the effectiveness of different AI tools could further refine our understanding of how technology can best be utilized to support personalized learning.

In conclusion, the integration of AI into the development of educational materials presents a promising avenue for enhancing language education, particularly in professional settings. By leveraging AI for the creation of customized learning resources, it seems that educators can significantly enhance the relevance and effectiveness of their teaching strategies, ultimately leading to better educational outcomes and more satisfying learning experiences for students.

This study provides a small contribution to the growing body of literature advocating for the increased use of technology in education and highlights the need for ongoing innovation and adaptation in teaching methodologies to meet the evolving needs of learners in the digital age.

Александар Кавгић

ЕФЕКТИ КОРИШЋЕЊА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ У НАСТАВИ
ЕНГЛЕСКОГ ЈЕЗИКА: СТУДИЈА СЛУЧАЈА НАСТАВЕ ЕНГЛЕСКОГ
У РАДНОМ ОКРУЖЕЊУ

Сажетак

Ова квантитативна пилот студија случаја у области наставе енглеског језика у радном окружењу (WE) имала је за циљ да утврди у којој мери употреба вештачке интелигенције (AI) у креирању наставних материјала може утицати на исходе учења и искуство студената. Студија је користила традиционалну експерименталну методологију са контролном и експерименталном групом којима су дати пре-тест и пост-тест за одређивање како степена језичких вештина тако и нивоа задовољства искуством учења. Експеримент је спроведен у малој ИТ компанији међу одраслим ученицима енглеског језика, укључујући појединце из три тимаса укупно 37 људи подељених у две групе. Контролна група од 18 људи је поучавана према стандардном уџбенику издавачке куће OUP (Business Results, комбинованом са International Edition of English for Customer Care) и додатним општим материјалима које је припремио наставник. Експериментална група од 19 људи похађала је наставу засновану искључиво на прилагођеним наставним материјале коју је створила вештачка интелигенција, којој је било задато да припреми материјале за учење прилагођене на основу компанијских имејлова, транскрипата телеконференцијских позива, објава на друштвеним мрежама и сличним материјала који су били претходно креирани или снимљене унутар компаније. AI је такође био задужен за припрему материјала за проверу знања (питања и квизови) који су били фокусирани на сваки појединачни део садржаја генерисаних на овај начин. Резултати експеримента показују да су и контролна и експериментална група побољшале језичке вештине које су биле фокус наставних материјала на статистички значајан начин, при чему контролна група није показала статистички значајна побољшања међу појединцима, док је експериментална група показала статистички значајна побољшања међу већином појединаца. Поред тога, експериментална група је известила о већем нивоу задовољства након завршетка курса. Укратко, резултати ове мале пилот студије указују на то да употреба AI у припреми ресурса за учење за наставу енглеског на радном месту доводи до бољег стицања релевантних вештина као и позитивнијег, субјективно мереног, искуства учења.

Кључне речи: AI, енглески језик на радном месту, наставни материјали, задовољство наставним процесом, настава енглеског на радном месту

BIBLIOGRAPHY

- Beevi, L. S., Vemuri, V. V., & Unnam, S. S. (2024). Sentiment analysis by using TextBlob and making this into a webpage. *AIP Conference Proceedings*, 2802(1). <https://pubs.aip.org/aip/acp/article/2802/1/120009/3126859>
- Dewi, H. K., Wardani, T. I., Rahim, N. A., Putri, R. E., & Pandin, M. G. R. (2021). *The use of AI (Artificial Intelligence) in English learning among university student: Case study in English Department, Universitas Airlangga*. <https://osf.io/preprints/sdnt/>
- Duarte, Fabio. (2024, March 30). *Number of ChatGPT Users (Apr 2024)* [Blog]. Exploding Topics. <https://explodingtopics.com/blog/chatgpt-users>
- Firat, M. (2023). *How chat GPT can transform autodidactic experiences and open education?* <https://osf.io/preprints/9ge8m/>
- Fitria, T. N. (2021). The use technology based on artificial intelligence in English teaching and learning. *ELT Echo: The Journal of English Language Teaching in Foreign Language Context*, 6(2), 213–223.
- Fitria, T. N. (2022). Using authentic material and created material (teacher-made) for English Language Teaching (ELT): Benefits and limitations. *JADEs Journal of Academia in English Education*, 3(2), 117–140.
- Jia, F., Sun, D., Ma, Q., & Looi, C. -K. (2022). Developing an AI-based learning system for L2 learners' authentic and ubiquitous learning in English language. *Sustainability*, 14(23), 15527.
- Jia, J. (2009). An AI framework to teach English as a foreign language: CSIEC. *Ai Magazine*, 30(2), 59–59.
- Lee, D., Kim, H., & Sung, S. -H. (2023). Development research on an AI English learning support system to facilitate learner-generated-context-based learning. *Educational Technology Research and Development*, 71(2), 629–666. <https://doi.org/10.1007/s11423-022-10172-2>
- Lightbown, P. M., & Spada, N. (2021). *How Languages Are Learned 5th Edition*. Oxford university press.
- Lucas, P. D. O., Maggioli, G. D., & López-Barrios, M. (2018). Teaching materials and their importance in ELT: Some perspectives from the Mercosur region. *The TESOL Encyclopedia of English Language Teaching*, 1–8.

- Marr, B. (2023, May 19). *A Short History Of ChatGPT: How We Got To Where We Are Today*. Forbes. <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/>
- Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, 16(1), 20, s13040-023-00339-9. <https://doi.org/10.1186/s13040-023-00339-9>
- Nguyen, M. H. (2013). EFL Students' Reflections on Peer Scaffolding in Making a Collaborative Oral Presentation. *English Language Teaching*, 6(4), 64–73.
- Nunan, D. (1988). The learner-centred curriculum: A study in second language teaching. (No Title). <https://cir.nii.ac.jp/crid/1130282272643475072>
- Nunan, D. (1999). *Second Language Teaching & Learning*. Heinle & Heinle Publishers. <https://eric.ed.gov/?id=ED441344>
- Opyr, M., Dobrovolska, S., & Panchyshyn, S. (2022). *Artificial intelligence for learning English*. <http://repository.pdmu.edu.ua/handle/123456789/20359>
- Peachey, N. (2013). A blended learning teacher development course for the development of blended learning in English Language Teaching. *Blended Learning in English Language Teaching: Course Design and Implementation*, 65.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge university press.
- Rukiati, E., Wicaksono, J. A., Taufan, G. T., & Suharsono, D. D. (2023). AI on learning English: Application, benefit, and threat. *Journal of Language, Communication, and Tourism*, 1(2), 32–40.
- Tomlinson, B. (2013). *Developing materials for language teaching*. Bloomsbury Publishing.
- Tomlinson, B., & Masuhara, H. (2013). Adult coursebooks. *ELT Journal*, 67(2), 233–249.

SOURCES

- Baade, K., Holloway, J., Scrivener, J., Turner, R., Davies, G., Shouler, A., Speck, C., & Wilden, S. (2016). *Business result: Advanced: student's book*. Oxford University Press.

Business result: Advanced: teacher's book (Second edition). (2022). Oxford University Press.

<http://www.nukat.edu.pl/nukat/icov/WALAZ/xx005635067.pdf>

Richey, R. (2007). *English for customer care*. Oxford University Press.