Pedagogika, 2025, 75(4), 401–422
http://pages.pedf.cuni.cz/pedagogika/

Research Paper

Using AI-Generated Songs in Foreign Language Learning

Tomáš Botlík Nuc, Marie Klusáčková

Abstract: *This study investigates the impact of AI-generated songs on vocabulary acquisition and learner motivation in foreign language education. Conducted at a secondary vocational school in Prague, the research involved learners of German as a second foreign language from the third-year classes. Using a quasi-experimental design with 60 observations – 30 in the experimental conditions and 30 in the control conditions – the study compared conventional vocabulary learning methods with an innovative approach involving the AI music platform Suno. In the experimental conditions, the learners achieved a significantly higher mean post-test score (6.2) compared to the control conditions (4.2), with a statistically significant difference confirmed at $p < 0.001$ and a large effect size (Cohen's $d > 0.8$). The survey data revealed that the learners preferred the song-based activity, citing increased engagement and enjoyment. These findings suggest that AI-generated songs can significantly enhance both vocabulary retention and student motivation, particularly in contexts requiring the acquisition of specialised terminology, which is often more challenging because of its complexity and the limited availability of learning resources.*

Keywords: *artificial intelligence, AI-generated songs, Suno, foreign language learning, vocabulary acquisition, learner motivation, secondary vocational education, specialised vocabulary*

Introduction

The use of Artificial Intelligence (AI) is becoming an integral part of education, and foreign language teaching and learning is no exception. AI brings new problems to the classroom, but also new possibilities. The possibilities that AI brings to education can be grouped into three categories: learner-facing, teacher-facing, and system-facing (Baker et al., 2019), which shows how broad its potential is. It can be a useful tool not only for learners, but also for teachers and school administrators. Many (if not all) major IT companies supplying software and hardware to schools use AI in their products. One example for all could be BETT 2025 (formerly known as the British Educational Training and Technology Show), where AI was at the centre of attention.

According to Pokrivčáková (2019), the main ways in which AI is used in foreign language education are: generating

DOI: 10.14712/23362189.2025.4919

personalised learning material; using machine translation tools; utilising AI writing assistants; conversing with chatbots applying AI-powered language learning software; relying on intelligent tutoring systems, and intelligent virtual reality. Using AI to create music or other creative content in the classroom is not explicitly among these main uses, but it could be seen as part of the first of these, generating personalised learning material. It might also be that Pokrivčáková did not see this use of AI in foreign language education as a major possibility, because song-generating AI platforms were not widely in use until about 2023–2024. Suno, the application the present study focuses on, was released in November 2024 and since then has continued to grow in popularity.

While AI might be seen as a danger to human creativity, Wu et al. (2021) argue that it is rather an empowering tool, making creativity more accessible and more inclusive than ever before and bringing new opportunities to various fields, education included. This type of attitude is one of the premises of this study. The Suno tool in particular offers the possibility of generating songs very quickly and efficiently, without the need to be a professional musician and having a whole band at one's disposal. Learners working with Suno can create music that aligns with their personal taste without being able to read sheet music or having deeper knowledge of song composition.

The use of songs in foreign language education is nothing new. One of the famous educators who have concerned themselves with music is Lozanov, who within the framework of suggestopedia recommended the use of music as a motivational tool that properly induces all thought and physical processes (Tepperwein, 1993, pp. 135–137). In suggestopedic methods, music is used to relax the mind of the learner. The results of Lozanov's authentic language teaching can be summarised as follows: beginner students assimilate an average of 80–90% of the given vocabulary (2,000 lexical units) on a receptive level during a 100-hour course. More than 60% of this vocabulary can be used actively in conversation. Furthermore, they are not afraid to speak with a foreigner. They communicate with errors, which, however, do not hinder understanding and they often have the desire to learn the foreign language further (ibid.).

Lozanov's findings show that music can have a positive influence on several aspects of language learning, namely vocabulary acquisition, learners' motivation, and confidence in communication. This study is interested in the first two. Even though the chosen method is very different from suggestopedia and is centred around the creation of music, it aims to explore the effect of music on vocabulary acquisition and learners' motivation too. Before the emergence of AI tools, it would have been more demanding and time-consuming to let learners create their own songs, but with AI, teachers now have the opportunity easily to let students create, learn and have fun at the same time. Seeing this potential, the authors of this study used Suno to help learners acquire

specialised vocabulary from the field of law and public administration, which is otherwise quite a challenging task. Specialised vocabulary is part of the compulsory curriculum at Czech secondary vocational schools but there are not many textbooks or additional resources available. AI tools have the potential to help fill this gap. This study aims to contribute practical insights for educators interested in leveraging AI-driven creative tasks in their own teaching practice.

The aim of this research is to explore the impact of AI-generated songs on both the linguistic and motivational outcomes of language learners, focusing therefore on the following two questions: **Q1** Does the integration of AI-generated songs improve vocabulary retention among secondary vocational school learners of German? **Q2** How do learners perceive the use of AI-generated songs in their language learning process? As the main tool, the AI-driven music creation platform called Suno was chosen, mainly because of it currently being the most used song-generating AI application.

To explore the impact of AI-generated songs on vocabulary retention, a quasi-experimental design with pre-test-post-test comparison was conducted. After the experiment, qualitative and quantitative feedback was collected in order to analyse the perceptions of learners.

Literature Review

When using AI-generated songs in foreign language learning, several aspects have to be taken into consideration. First, there is the role AI plays in education, then, more specifically, the potential both AI and music bring to foreign language education.

AI in Education: Potential and Dangers

Artificial intelligence (AI) has the potential to revolutionise the way we learn and teach (Kamalov et al., 2023). AI can help with personalisation of the learning experience for each student, allowing them to learn at their own pace and according to their individual needs and abilities and reflecting their expectations. It can also be a way to increase student engagement (Harry, 2023).

The survey of National AI Policy Strategies conducted by Schiff (2022) has shown, however, that the use of AI in education still has many limits and is generally oriented towards education for AI (creating AI experts) and not AI for education. Policymakers might be in danger of missing the transformative potential AI has for education and also sufficiently recognising its ethical implications, while the number of ethics incidents related to AI is growing rapidly.

The ethical issues related to AI in education include the following: data privacy and security, bias and discrimination, plagiarism and academic integrity, and the teacher-student relationship (Kamalov et al., 2023). When using AI in education, it is necessary to ensure students' data is safe, the potential biases are addressed,

students are informed about plagiarism, and the well-being and human connection of all participants is cared for. These guidelines are further specified and partly extended by institutions such as the European Union (for more details see European Union, 2022). While these dangers might be of lesser importance when it comes to the way AI is used in this study, it would not be wise to forget about the fact that AI has its limits. When creating songs for classroom use, users' personal data is not required and there is little potential for discrimination. That having been said, a debate about the way original art is used to train AI models might bring interesting insights.

AI-generated Content and Vocabulary Acquisition

Several studies across different age groups and learning contexts have compared the effect on vocabulary acquisition of conventional methods with that of AI-enhanced materials. While most of the conventional methods used for vocabulary teaching are effective, AI has the potential to enhance the effect and individualise the materials. Agnes and Srinivasan (2024) used AI-generated mnemonics to enhance Anki flashcards. While both their groups of intermediate-level English learners displayed improvement in vocabulary retention, the group using AI-generated mnemonics showed a more pronounced enhancement compared to the group using flashcards without AI-generated mnemonics. Another study showed that AI-generated stories can help preschool children improve their receptive and expressive vocabulary skills to the same extent that traditional stories can (Cho & Yim, 2024). Their findings suggest that AI can be used to personalise learning materials in a very time-efficient manner, even when it comes to creative materials such as pictures, stories, or songs. Wei-Xun and Jia-Ying (2024) support such findings when they advocate the use of AI-driven language learning apps in the EFL classroom. In their research, Wei-Xun and Jia-Ying compared the results of an experimental group using AI-driven applications to the results of a control group following conventional instruction based on a textbook. They found that AI-based intervention had a significant positive effect on both vocabulary acquisition and motivation levels among students.

Whilst the above-mentioned studies provide only episodic evidence for the potential of AI tools for the area of vocabulary acquisition and retention, systematic reviews are also emerging. Yang (2025) states that AI tools "could significantly improve vocabulary acquisition among K-12 learners". Yang analysed 30 empirical studies and found that although there has been a rapid increase in research in this field (with an ongoing upward trend), a unified theoretical framework is still lacking. Another conclusion drawn from the review is that research methods in this area primarily employ a mixed-methods design approach, combining quantitative and qualitative methods. This study

attempts to follow this trend and uses a mixed-methods design approach to provide deeper insights into the potential of AI-generated songs for vocabulary acquisition.

Songs in Foreign Language Education

After outlining the broader potential of AI in education, this section turns to the specific role of music in language learning. Using music to enhance language education is nothing new, yet it is still viewed as something innovative and often not used to its full potential (Pokorný, 2021). While "teachers seem to be positive about the use of music in the FL classroom, the incorporation appears rather occasional" (Degrave, 2019, p. 412), often because of the lack of proper resources and the difficulties connected with searching for suitable songs for the topics covered (Tse, 2015). As this study attempts to demonstrate, AI music could help overcome this difficulty, as teachers no longer have to look for suitable songs, but can create them themselves or even let the learners join the creative process as part of learning.

The specific ways in which music can be used to support language learning are diverse. Degrave (2019) classifies them into three main categories: the use of music without lyrics (background instrumental music), the use of songs, and the use of rhythmic activities. A study conducted by de Groot (2006) found that background music increased vocabulary retention by 5.8% when learning frequent words and by 11.6% when learning infrequent words, which suggests that music can help with the acquisition of more difficult lexical items such as specialised vocabulary. The results, however, varied across learners, calling for a deeper analysis of factors such as personal music preference.

There is a lot of research supporting the benefits of the use of songs (with lyrics) in foreign language education, both for memory and engagement. Israel (2013) struggled to motivate learners from various backgrounds for the poetry analysis which is part of their ESL curriculum. She found that turning to music first brought surprisingly high engagement to the subsequent lessons. Apart from motivation, the study also found that using contemporary songs helped learners develop analytical skills, improved written and oral communication, and fostered a deeper connection with the target language, leading to increased confidence and academic success.

Ludke (2010) devoted her work to listening to songs and singing for learning modern foreign languages (L2). This research (based on three sub-research studies) supports the claim that singing aids L2 learning, showing that even beginners benefit as much from songs as from spoken materials, with improvements in listening, speaking, vocabulary, and grammar. Teachers should integrate songs into the curriculum regularly rather than using them as oc-

casional rewards, as consistent musical activities can enhance both language acquisition and student engagement. The study by Schunk (1999) supports this conclusion too: children who heard an intervention text sung had a significantly higher vocabulary gain over children who only heard it spoken. They averaged 6.5 words learned, as opposed to the text group's 2.5 words. Clearly, song and music are effective language teaching tools. Stansell (2005) states that working with music allows for a more playful approach to language, which can promote a communicative approach to language learning that involves social interaction, group work, and peer discussion. Music provides relief from vocabulary tests and grammar drills. The use of songs and music today offers a more natural and holistic way of teaching language. Overall, these studies support the use of songs as a complementary strategy in foreign language education, aiding both vocabulary acquisition and motivation.

Previous research, as quoted above, has shown that AI-generated content and music can both have a positive effect on vocabulary retention and learner motivation. On the basis of these findings, the following hypotheses were formulated:

H1 Learners who create AI-generated songs will demonstrate higher vocabulary retention compared to those who engage in conventional learning methods.

H2 Learners using AI-generated songs will report higher levels of motivation than those learning without them.

Methodology

Building on the insights derived from the literature, the following sections outline the methodology employed to investigate the hypotheses that were formulated.

Research Design

To explore the hypotheses, this study employs an action research approach with quasi-experimental features. Two forms of analysis were conducted: a quantitative comparison and a questionnaire. The comparison used pre-test and post-test results to explore vocabulary acquisition and the effectiveness of the instruction method. To ensure comparability, both the experimental and control groups completed the same vocabulary test at the beginning and end of the procedure. The ten-item test assessed learners' ability to recall German words from their Czech translations and required about five to ten minutes to complete. A complete list of the pre- and post-test items is provided in Appendix A. The questionnaire was in a mixed format, combining quantitative (Likert-scale) and qualitative (open-ended) items to evaluate the learners' motivation and perceptions of the activity (Rod, 2012).

The research took place at a secondary vocational school in Prague, within the context of third-year German language classes (roughly equivalent to Grade 12 in the American system or Year 13 in the British system). Three groups of learners from the same field of study participated in the research. The participants were not

randomly assigned to the experimental and control groups, as they were pre-existing classes. Therefore, to minimise potential group-related bias, the study was conducted in two rounds:

- in round 1, two groups worked in experimental conditions (using AI-generated songs via Suno), while the third group served as the control group (traditional vocabulary instruction without music or AI);
- in round 2, the learners worked with another set of words. The roles were reversed: the original control group received experimental instruction, and the two original experimental groups served as the control groups.

This design allowed for between-group comparisons while respecting the pre-existing classroom structure. While two groups might have been enough to explore the effect of the conventional versus experimental instruction, three groups were chosen to increase the sample and also bring more diversity to the experiment, as they are normally taught by different teachers and therefore used to different styles of teaching.

Research Process

The data collection occurred from February to April 2025. The aim was to compare the effectiveness of AI-generated songs with traditional vocabulary instruction in a secondary vocational school setting.

The process involved several key phases: pre-test administration, classroom instruction and activities (conventional/

Table 1. Research Process, Round 1 (Topic 1)

Topic 1	Group 1	Group 2	Group 3
Step 1	pre-test		
Step 2	conventional instruction	**experimental** instruction	**experimental** instruction
Step 3	post-test questionnaire		
Step 4	two-sample t-test, Cohen's *d* survey data analysis		

Table 2. Research Process, Round 2 (Topic 2)

Topic 2	Group 1	Group 2	Group 3
Step 1	pre-test		
Step 2	**experimental** instruction	conventional instruction	conventional instruction
Step 3	post-test questionnaire		
Step 4	two-sample t-test, Cohen's *d* survey data analysis		

experimental), post-test and questionnaire administration, and data analysis (Tables 1 and 2). A more detailed description of the intervention is provided below.

Intervention procedure –control conditions

Lesson 1: The learners completed a pre-test, followed by collective translation and review of the individual words. Following this, they were tasked with using the words in context by writing sentences or short texts. They were allowed to use a dictionary (but not translators or AI tools). The teacher reviewed their texts and corrected them when needed.

Lesson 2 (two days after the pre-test): The learners were asked to read two texts – their own and one produced by a classmate. Then they individually completed a set of exercises focusing on the target vocabulary.

Lesson 3 (one week after the pre-test): The learners completed a post-test and then participated in a competition for the best text. As the last step, they filled in a questionnaire focused on the evaluation of the activity by the learners.

Intervention procedure – experimental conditions

Lesson 1: The learners completed a pre-test, followed by collective translation and review of the individual words. Following this, they were tasked with using the words in context by writing sentences or short texts. They were allowed to use a dictionary (but not translators or AI tools). After teacher review of the texts, the learners were asked to copy their texts into Suno and generate songs, using genre descriptors such as 'pop', 'rap', 'lullaby', or 'nostalgic' to tailor the AI-generated songs to their taste.

Lesson 2 (two days after the pre-test): The learners were asked to listen to two songs – their own and one produced by a classmate. Then they individually completed a set of exercises focusing on the target vocabulary. Examples of the texts written by the learners and the exercises completed can be found in Appendices B and C respectively.

Lesson 3 (one week after the pre-test): The learners completed a post-test and then participated in a competition for the best song. As the last step, they filled in a questionnaire focused on the evaluation of the activity by the learners.

Context

The German language was selected for a number of reasons. The proficiency levels of the learners within the respective language groups are well-balanced. There is a notable absence of significant disparities amongst the learners, a phenomenon that is not always observed in the teaching of other languages, such as English. Additionally, the vast majority of the learners do not engage with the language outside the school environment. Finally, the secondary vocational school was selected because of the Framework Educational Programme (RVP, 2023) stipu-

lating that schools must offer instruction in a specialised foreign language in which learners have minimal prior experience. Consequently, it can be deduced that all the knowledge and skills acquired in this respect take place exclusively in the context of school teaching.

The RVP for vocational education requieres in terms of the acquisition of a vocabulary of at least 2,300 lexical units per course of study, of which specialised terminology accounts for at least 20% at B1 level and 15% at A2 level (RVP, 2023). A2 is the minimum exit level for a second foreign language, but schools may individually require a higher level. The school where this research was conducted requires A2 for a second foreign language; the figure is therefore 15%. However, it is assumed that a similar approach to vocabulary teaching might be used for general vocabulary as well, not only for specialised vocabulary.

Participants

The research investigation was conducted at a secondary vocational school in Prague, within the context of German language classes with learners in the third grade. These learners had been learning German as a second foreign language since Year 1. Three groups of learners from the same subject area were selected for the research study, with each group having a quota of four German lessons per week.

The learners' language proficiency levels were determined using the Common European Framework of Reference for Languages, with the teachers administering indicative placement tests in January 2025. The results of these tests indicated that the learners' language proficiency levels ranged from A1 to A2. The age of the participants was in the range of 17 to 18 years old. During the questionnaire survey, it was found that almost all the participants involved had no previous experience with the Suno tool.

Data Collection and Analysis

A total of 34 students participated in the study, contributing 60 post-test scores across two instructional conditions. Since some participants completed only one condition (control or experimental), the data was treated as independent, and a two-sample independent t-test was used to compare the post-test results between the two groups.

From the first group of learners 15 results were included in the analysis, from the second group 27 results, and from the third group 18 results. Although the groups were originally of similar size (14–15 learners in each), the number of results differs. This is caused by the fact that the study took place during times with high sickness rates and out of the 44 learners enrolled, ten missed some or all the steps in round 1 or round 2 of the research process and were therefore not included in the analysis.

Furthermore, the results of the questionnaire surveys were documented and compared. Additionally, the learners' lyrics or their musicalised songs were ob-

tained from the Suno app but not analysed further, as the goal of this research does not include a detailed linguistic analysis of the learners' production.

The Belgian platform BookWidgets was utilised for the implementation of the pre-tests, post-tests, and interactive exercises. The Google Forms tool was utilised for the administration of the questionnaire surveys.

Ethical Considerations

Regarding the ethical issues of the study, the participants in this study were assured of their voluntary participation, anonymity, confidentiality, and the right to refuse to answer and/or withdraw from the study at any stage. They provided their informed consent to participation in the study in writing. For underage learners, informed consent was also secured from their legal guardians following a clear explanation of the study's aims and methods.

Results

Quantitative Results

Descriptive statistics: In the pre-tests, the learners scored either 0 or 1 in both the control and experimental conditions. This shows that before the experiment, they were not familiar with the chosen vocabulary items and therefore the results could not have been influenced by differ-

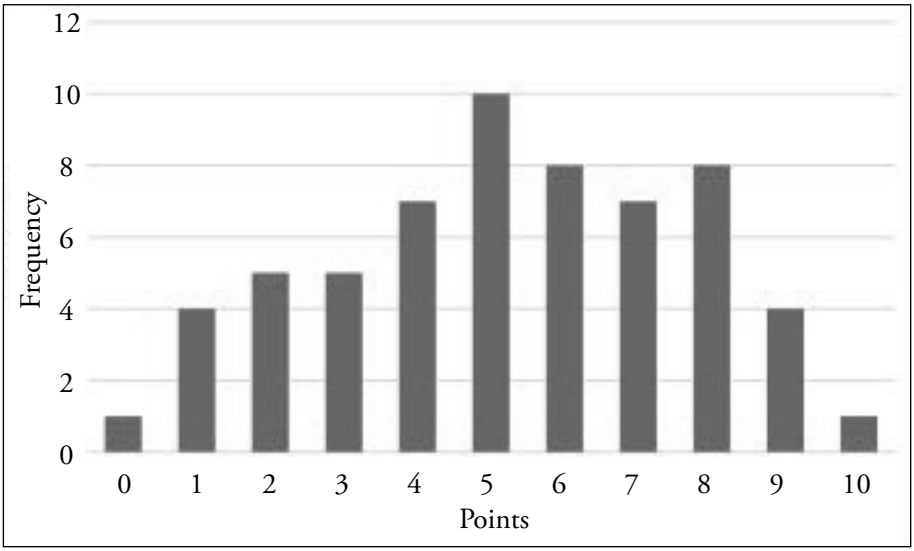

Figure 1. Histogram of the Post-Test Results. N (post-test results) = 60 (includes results from both the experimental and control conditions)

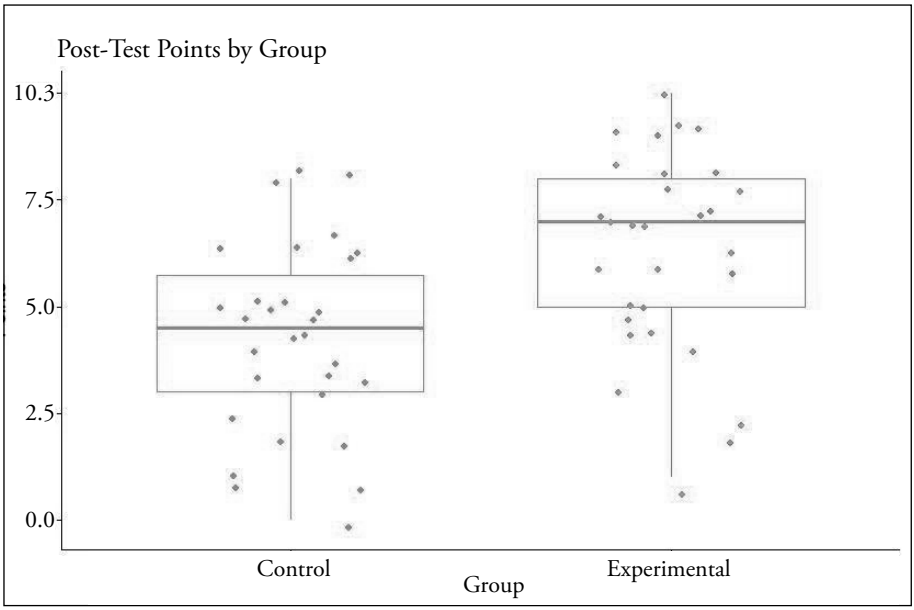

Figure 2. Post-Test Points by Group. N (post-test results) = 60

ent levels of familiarity among the participants. As shown in Figure 1, the distribution of the post-test scores tends towards normality, with the most frequent result being five points. This indicates a relatively even spread of performance levels among the learners after the intervention.

Figure 2 presents a boxplot comparison of the post-test scores between the control and experimental conditions. The thicker line within each box denotes the median, while the box itself captures the interquartile range (IQR), showing the middle 50%. The experimental post-test results' median score (7) notably surpasses that of the results collected in control conditions (4.5), indicating a notable improvement in vocabulary retention.

The results were the same when the mean was used instead of the median. The mean post-test result from the experimental conditions was 6.2 points, two points higher than from the control conditions (see Table 3). The difference between the experimental and control conditions in the mean post-test result was statistically significant at the 0.05 level and $t(58) = 3.436$. The t-value of 3.436 reflects the ratio between the observed difference in the group means and the variability within the data (Duncan, 1975). A higher t-value indicates a greater likelihood that the difference

Table 3. T-test Comparison of Post-Test Points by Group

	Control (N = 30)	Experimental (N = 30)	t-value	*p*
Mean	4.23	6.23	3.436	0.001
SD	2.18	2.33		

between the groups is not due to random chance. In this case, the result exceeds the critical value for 58 degrees of freedom at the $p < 0.05$ level, confirming that the observed effect is statistically significant. This confirms the first hypothesis, namely that the learners working with the AI tool would remember more words than those who did not use the AI tool.

In order to determine the size of the difference between the performance of the learners in the experimental and control conditions, Cohen's *d* was calculated. The result of $d = 0.887$ indicates that the difference in the improvement in the learners' performance when the experimental instruction is applied is large. The Cohen's *d* coefficient can be interpreted as the percentage of people in one group who outnumber the average member of the other group. For individual values of Cohen's *d*, this percentage can be accurately determined using the values of the normal distribution function (Soukup, 2013).

According to Cohen (2013), $d = 0.2$ corresponds to a small effect; $d = 0.5$ to a medium effect; $d \geq 0.8$ to a large effect. Consequently, it can be concluded that the experimental teaching method had a significantly positive impact on the students' performance in comparison to the conventional approach.

Questionnaire

A questionnaire survey was distributed to the learners after both rounds of the experiment. The questions for the control conditions and experimental conditions differed slightly with regard to the characteristics of the specific activities, but not in the main areas – the enjoyment and perceived usefulness of the whole process. The answers were gathered from all 34 participants, both in the control conditions and in the experimental ones, to allow for comparison.

The analysis of the results revealed two key areas of interest: the satisfaction of the learners with the method for acquiring vocabulary, and their subjective perception of the effectiveness of that method.

Regarding the first area of interest, satisfaction, in the experimental conditions, the learners demonstrated greater enjoyment of the vocabulary acquisition process (Figure 3). When asked about the individual activities in the questionnaire administered after the experimental instruction, the majority of the learners indicated that setting the text to music using Suno was the most enjoyable aspect of the process. The second most appreciated activity was listening to the songs

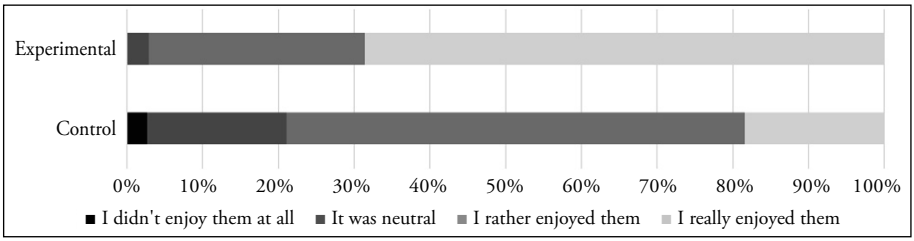

Figure 3. Enjoyment of the Activities

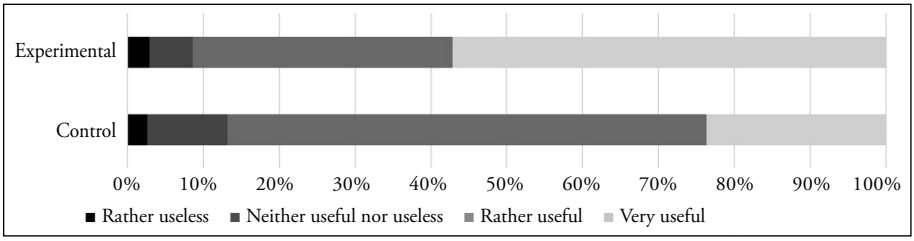

Figure 4. Usefulness of the Activities

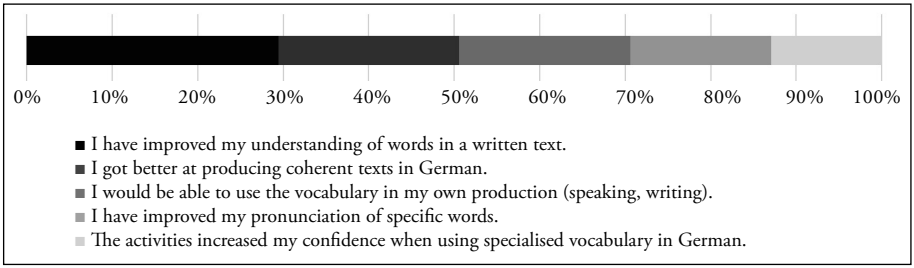

Figure 5. Statements Indicating Skills Learned – Control Conditions

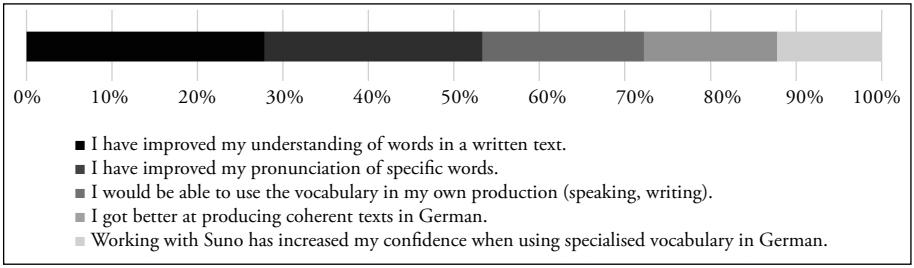

Figure 6. Statements Indicating Skills Learned – Experimental Conditions

created by themselves and their peers. The competitive element (voting for the best song) was mentioned only marginally, in sharp contrast to the answers in the control conditions, where most learners cited the competition as the most enjoyable part of the process. It appears that the innovative use of the AI tool overshadowed the motivational effect of the game, which tends to be the most engaging element in conventional instruction. Notably, after the experimental instruction, none of the learners expressed a negative opinion about working with the tool. The results overall are in favour of the second hypothesis of this research, namely that learners using AI-generated songs will report higher levels of motivation than those learning without them.

As for the perceived effectiveness, it is evident that the learners derived greater utility from the experimental instruction (Figure 4). However, it is important to note that the same number of learners rated this process as ineffective in both groups.

After evaluating the overall usefulness of the process, the participants were also asked which skills they believed they had improved the most. In the control conditions (Figure 5), the greatest improvements were reported in the following areas: comprehension of words within a text (1), the ability to produce coherent written text (2), and the use of vocabulary in their own written or spoken production (3). In the experimental conditions (Figure 6), two categories were repeated, namely understanding the vocabulary in context (1) and using the vocabulary in one's own production (3). However, the second place was taken by improvement of pronunciation (2).

The last part of the questionnaire offered space for additional comments and observations. Several comments from learners who received experimental instruction in the first phase of testing and conventional instruction in the second phase indicated that next time they would prefer an activity where they work with the Suno tool.

The learners were then invited to share their reflections on the experience, specifically addressing the questions: "What did you enjoy about it? What aspects of these activities did you not find enjoyable? What factors might have contributed to enhanced learning outcomes?" Any further comments were also invited.

Most of the learners expressed a preference for undertaking a similar or identical activity involving songs in the future. It was further noted that all the participants in the study concurred that the process of songwriting was more suitable for their needs than just writing text, with the respondents calling the experimental activities "more enjoyable, varied, effective, and interesting".

When asked to suggest ways in which learning could be enhanced, only a small number of learners offered any contributions. The following suggestions were made: the incorporation of additional activities to practise the vocabulary more, then also adding more games, and potentially bingo to assist with vocabulary de-

velopment. The suggestion was also made to practise the vocabulary in a manner similar to Dobble in German, given the necessity for rapid reaction time, connected to fast recall skills.

In response to a query regarding further remarks, two learners articulated that the second topic posed a greater challenge. "The second topic was more difficult, but I am uncertain whether this was due to the vocabulary being more challenging or whether it was simply not made more difficult." Four additional learners indicated that they found the second text more challenging to write because the words were not as closely related.

Conclusion

Limits of Research

This study has several limitations that should be acknowledged. One of the limitations may be the fact that the vocabulary used during the first and second phases was different. Despite the fact that it was specialised vocabulary, the second topic (as some learners in the focus groups indicated) may have been more difficult to grasp for text production.

A second limitation lies in the relatively small sample size. While the findings offer promising insights into the use of the generative AI tool Suno in vocational language teaching, further research with a larger and more diverse sample is necessary to confirm the generalisability of the results.

Another limitation of the research is the filtering of the participants. Several participants had to be excluded from the data analysis because they did not participate in all the phases of the research because of their absence from school. Additionally, the results may be somewhat biased by the order of the two types of instruction. The groups that started with the experimental instruction could have been less motivated when they had to work without the AI tool in the second round of the experiment. However, only one participant stated in the questionnaire that they had not enjoyed the activities in the control conditions at all. Still, the direction of change between the control and experimental instruction could be a topic for further analysis. In the future consideration can be given as to whether to conduct similar research with the same group, which will become both the experimental and control group.

For future research, it would be beneficial to investigate whether similar results could be achieved through long-term repetition of the methods used. In other words, it would be important to determine whether the observed effects were due to the novelty of the AI tool or whether sustained engagement with the songs would continue to support vocabulary acquisition and learner motivation over time. Such research would help control for the novelty effect, which, as Lee (2011) notes, can temporarily boost learner satisfaction and thus influence learning outcomes.

Conclusion and Implications for Teaching

The findings of this study demonstrate that incorporating AI-generated songs into foreign language instruction can significantly enhance vocabulary acquisition and learner motivation. When working with the Suno tool, the learners not only outperformed their peers in vocabulary retention tests but also reported higher levels of enjoyment and engagement during the learning process. Both research hypotheses were thus supported. This aligns with previous research, such as Solehudin et al. (2024), who found that AI-generated songs created by Suno significantly improved the motivation, listening skills, and vocabulary acquisition of 16–18-year-old learners learning foreign languages.

The ability to personalise and engage with content creatively through AI-generated music appears to be particularly effective for teaching specialised vocabulary — an area often underserved by traditional materials, especially in vocational education contexts. While previous studies, as noted in the introduction, have explored the use of songs and AI tools in language learning, few have specifically examined the impact of student-generated AI-assisted songs on learning outcomes in vocational education. This study offers a novel perspective that may contribute to emerging discussions around the pedagogical potential of creative AI tools in this context.

Although this study faced limitations, such as a relatively small sample size and varying difficulty of the vocabulary topics, the results are promising. They suggest that integrating creative AI tools into the language classroom can offer a meaningful and motivating learning experience. The enthusiastic response from learners further indicates strong potential for the broader adoption of such tools in educational settings.

Future research could expand this approach with larger and more diverse participant groups and examine the long-term effects on vocabulary retention and broader language skills. Nevertheless, this study provides valuable insights for educators seeking innovative technology-enhanced methods to make language learning more effective, personalised, and engaging.

During the investigation, some technical issues were encountered, pointing to the limits AI tools have. It was discovered that when a group of learners sought to register for Suno at the same time, the system perceived this as a threat and only permitted registration of the first few. It is therefore recommended that registrations for Suno be conducted in a continuous manner prior to the official implementation of the system.

Because the results of this research show improvement not only in receptive skills such as comprehension of vocabulary, but also in productive skills, it is recommended to use Suno and similar AI tools to foster both areas. While the pronunciation used by AI platforms might not be as accu-

rate as that of a real person (native speaker), it is improving constantly and can be useful for pronunciation practice. Apart from pronunciation, writing and speaking can be part of AI-based activities as well. The flexibility of the tool enables easy integration of different target skills.

The motivational effect of AI tools shown by the results of this research and previous studies might prove particularly important in areas or subjects that are less popular among learners. German language education, and particularly specialised German vocabulary, might be one such area. The teachers involved in this research reported that learners at secondary vocational schools typically do not consider German their favourite subject, often on the basis of previous experience. However, creative AI activities (such as lyrics and music composition) show the potential to raise learners' engagement and it is therefore recommended that teachers experiment with them, bringing playfulness, creativity, personalisation, and innovation to classes of any subject, not only languages. During the research, several learners themselves reported that they turned other study texts into music to make it easier for them to learn other subjects. Before introducing new digital tools to their classes, nevertheless, it is advisable to prepare adequately and test them, for example among teachers.

References

Agnes, D., & Srinivasan, R. (2024). Fostering vocabulary memorization: Exploring the impact of AI-Generated mnemonic keywords on vocabulary learning through Anki flashcards. *World Journal of English Language, 14*(2).

Baker, T., Smith, L. & Anissa, N. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. *Nesta*. https://www.nesta.org.uk/report/education-rebooted/

Cho, J. E., & Yim, D. (2024). The impact of personalized book reading activities using AI-generated stories on vocabulary acquisition in preschool children. *Communication Sciences and Disorders*, *29*(4), 686–702.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Routledge.

Degrave, P. (2019). Music in the foreign language classroom: How and why. *Journal of Language Teaching and Research*, *10*(3), 412–420.

Duncan, D. B. (1975). T tests and intervals for comparisons suggested by the data. *Biometrics*, *31*(2), 339–359.

European Union. (2022). *Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators*. https://education.ec.europa.eu/

de Groot, A. M. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning*, *56*(3), 463–506.

Harry, A. (2023). Role of AI in education. *Interdiciplinary Journal & Hummanity (INJURITY)*, *2*(3).

Israel, H. F. (2013). Language learning enhanced by music and song. *Literacy Information and Computer Education Journal*, *2*(1), 1360–1366.

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, *15*(16), 12451.

Lee, Y. J. (2011). A study on the effect of teaching innovation on learning effectiveness with learning satisfaction as a mediator. *World Transactions on Engineering and Technology Education*, *9*(2), 92–101.

Ludke, K. M. (2010). *Songs and singing in foreign language learning.* (Doctoral dissertation). Pepperdine University, The University of Edinburgh.

Pokorný, P. (2021). Využití písní v hodinách cizích jazyků se zaměřením na výuku německého jazyka. *Cizí jazyky, 4(*64).

Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education.*

Rod, A. (2012). Likertovo škálování. *E-Logos Electronic Journal for Philosophy*, *13*, 2–14.

RVP. (2023). *Rámcový vzdělávací program pro obor vzdělání 68 – 43 – M/01 Veřejnosprávní činnost.* Praha: MŠMT, 2023. https://www.edu.cz/wp-content/uploads/2020/08/68-43-M01_Verejnospravni_cinnost_2020_zari_rev.pdf

Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*, *32*(3), 527–563.

Schunk, H. A. (1999). The effect of singing paired with signing on receptive vocabulary skills of elementary ESL students. *Journal of Music Therapy, 36,* 110–124.

Stansell, J. W. (2005). The use of music for learning languages: A review of the literature. University of Illinois at Urbana-Champaign. *Literacy Information and Computer Education Journal, Special 2*(1), 1360–1366.

Solehudin, M., Hanifansyah, N., & Izzuddin, I. F. (2024). Enhancing Arabic listening and vocabulary acquisition through AI-powered music: A study on suno AI. *Arabiyat: Journal of Arabic Education & Arabic Studies*, *11*(2).

Soukup, P. (2013). Věcná významnost výsledků a její možnosti měření. *Data & Research: SDA Info*, *7*(2).

Tepperwein, K. (1993). *Umění lehce se učit.* Stanovum.

Tse, A. Y. (2015). Malaysian teachers' perspectives on using songs in English language teaching. *International Journal of Social Science and Humanity, 5*(1), 87–89.

Wei-Xun, L., & Jia-Ying, Z. (2024). Impact of AI-Driven language learning apps on vocabulary acquisition among English learners. *Research Studies in English Language Teaching and Learning, 2*(1), 2–10.

Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., & Shidujaman, M. (2021). AI creativity and the human-AI co-creation model. In M. Kurosu, M. (Ed.), *Human-computer interaction: Theory, methods and tools (HCII 2021)* (pp. 171–190). Springer.

Yang, Y. (2025). AI-supported L2 vocabulary acquisition – a systematic review from 2015 to 2023. *Education and Information Technologies*, *30*(13), 1–35.

Tomáš Botlík Nuc
Charles University, Faculty of Education, Department of German Studies;
email: tomas.botlik@pedf.cuni.cz

Marie Klusáčková
Charles University, Faculty of Education, Department of German Studies

BOTLÍK NUC, T., KLUSÁČKOVÁ, M. Využití písní generovaných umělou inteligencí ve výuce cizích jazyků

Tato výzkumná studie zkoumá dopad písní generovaných umělou inteligencí na osvojování slovní zásoby a na motivaci žáků v oblasti výuky cizích jazyků. Výzkum byl realizován na střední odborné škole v Praze a zúčastnili se ho žáci třetích ročníků, kteří se učí němčinu jako druhý cizí jazyk. Pomocí kvaziexperimentálního designu s celkem 60 pozorováními — 30 v experimentálních podmínkách a 30 v kontrolních podmínkách — byla porovnávána konvenční metoda učení se slovní zásoby s inovativním přístupem využívajícím platformu Suno. Platforma Suno je založená na umělé inteligenci a určená ke generování písní. V experimentálních podmínkách dosáhli žáci výrazně vyššího průměrného skóre v post-testu (6,2) oproti kontrolním podmínkám (4,2), přičemž rozdíl byl statisticky významný na hladině $p < 0{,}001$ a měl velký efekt (Cohenovo $d > 0{,}8$). Dotazníkové šetření ukázalo, že žáci upřednostňovali výuku prostřednictvím zhudebňování textů, přičemž také zmiňovali vyšší míru vlastního zapojení a zábavnosti u využití nástroje Suno pro výuku. Zjištění naznačují, že písně generované umělou inteligencí mohou výrazně podpořit jak zapamatování slovní zásoby, tak motivaci žáků, zejména v kontextech, které vyžadují osvojení odborné terminologie, jež bývá náročnější kvůli své složitosti a omezeným výukovým zdrojům.

Klíčová slova: *umělá inteligence, písně generované umělou inteligencí, Suno, výuka cizích jazyků, osvojování slovní zásoby, motivace žáků, střední odborné vzdělávání, odborná slovní zásoba*

Appendix

Appendix A: Vocabulary items used in the pre-/post-tests

Pre-/post-test 1: vocabulary connected to contracts (specifically rental agreements)

- der Mietvertrag
- der Mieter
- der Vermieter
- wohnhaft in
- die Kosten
- sich verpflichten
- die Hausordnung
- die Wohnfläche
- auf unbestimmte Zeit
- kündigen

Pre-/post-test 2: vocabulary connected to politics

- die Verfassung
- die Regierung
- das Staatsoberhaupt
- der Regierungschef
- das Höchstgericht
- leiten
- wählen
- vorschlagen
- vertreten
- die Freiheit

Commentary: Both lists contain a mix of vocabulary items related to topics that form part of the participants' specialised curriculum and are required, among other things, for the Maturita exam. While the majority of the items are nouns, other parts of speech were also included to ensure that the vocabulary would be applicable and useful in real-life situations.

Appendix B: Examples of texts written by students

Example 1:

Ich bin ein **Mieter** mit einem **Mietvertrag** mit meinem **Vermieter**. Die **Kosten** sind sehr groß und ich habe kein Geld. Der Vermieter ist **wohnhaft** in Prag. Ich habe mich für 4 Jahre **verpflichtet**. Die **Hausordnung** ist schrecklich, aber ich kann nicht **kündigen**. Der einzige Grund, warum ich hier wohne, ist die **Wohnfläche**. Aber ich will nie einen Mietvertrag **auf unbestimmte Zeit** haben.

Example 2:

Der **Regierungschef** und das **Staatsoberhaupt** haben die **Verfassung vorgeschlagen**. Aber die Verfassung gibt keine **Freiheit**, so haben Leute eine neue **Regierung gewählt**, um sich besser **vertreten** zu können. Jetzt **leiten** die neue Regierung und das **Höchstgericht** einen Krieg gegen das Staatsoberhaupt und den Regierungschef.

Appendix C: Examples of exercises used to practice target vocabulary in context

Exercise 1:

Was fehlt?

Mieter Mietvertrag kündigen Wohnfläche Die Kosten sind hoch, aber ich brauche große _____.

Mein _____ ist auf unbestimmte Zeit aber mein Vermieter möchte meinen Vertrag _____. Das macht keinen Sinn, weil ich ein guter _____ bin.

Exercise 2:

Was fehlt?

Höchstgericht Freiheit wählen Verfassung Staatsoberhaupt Regierung leitet vorschlagen Präsident vorschlägt Die _____ ist sehr wichtig für ein Land, weil sie die Regeln für die _____ und die Rechte der Menschen bestimmt. Das _____ vertritt das Land in anderen Ländern und spricht mit anderen Politikern. Der Regierungschef _____ die Regierung und arbeitet mit seinen Ministern zusammen.

Die Bürger _____ ihre Vertreter, die neue Gesetze _____ und über wichtige Themen entscheiden. Das _____ prüft, ob die Gesetze richtig sind und schützt die _____ der Menschen. So funktioniert die Politik in einem demokratischen Land.

Appendix D: Excerpts from the questionnaire

The responses to these questions were used to generate the graphs presented in the analysis. In instances where one or more response options were not selected by any participant, those responses were excluded from the graphs.

Jak tě aktivity (psaní textu, procvičovací cvičení, soutěž) zaměřené na odbornou slovní zásobu bavily?

☐ Velmi mě bavily
☐ Spíše mě bavily
☐ Byly neutrální
☐ Spíše mě nebavily
☐ Vůbec mě nebavily

Jak užitečné ti aktivity připadaly k naučení se nové slovní zásoby?

☐ Velmi užitečné
☐ Spíše užitečné
☐ Ani užitečné ani neužitečné
☐ Spíše neužitečné
☐ Rozhodně neužitečné

Statements – experimental group

Označ všechna tvrzení, která pro tebe platí:

☐ Zlepšil/a jsem se v porozumění slovíčkům v textu (např. smlouvě, inzerátu).
☐ Dokázal bych slovíčka použít při vlastní produkci (mluvení, psaní textu).
☐ Práce se Suno zvýšila moji sebedůvěru při používání odborné slovní zásoby v němčině.
☐ Zlepšil/a jsem si výslovnost konkrétních slovíček.
☐ Zlepšil jsem se při tvorbě souvislého textu v němčině.
☐ Žádné tvrzení pro mě neplatí.

Statements – control group

Označ všechna tvrzení, která pro tebe platí:

☐ Zlepšil/a jsem se v porozumění slovíčkům v textu (např. smlouvě, inzerátu).
☐ Dokázal bych slovíčka použít při vlastní produkci (mluvení, psaní textu).

- ☐ Aktivity zvýšily moji sebedůvěru při používání odborné slovní zásoby v němčině.
- ☐ Zlepšil/a jsem si výslovnost konkrétních slovíček.
- ☐ Zlepšil jsem se při tvorbě souvislého textu v němčině.
- ☐ Žádné tvrzení pro mě neplatí.