



The Effect of LyricsTraining on Listening Achievement and Listening Anxiety in Foreign Language Education¹

ARTICLE TYPE	Received Date	Accepted Date	Published Date
Research Article	03.05.2025	06.30.2025	08.16.2025

Uğur Genç ²

Ministry of National Education

Tuba Acar Erdol ³

Ordu University

Abstract

The aim of this study is to examine the effect of LyricsTraining on the English listening achievement and listening anxiety of 7th grade middle school students. The research was conducted during the first semester of the 2023–2024 academic year at a public middle school located in a mid-sized city in Türkiye. A quasi-experimental design involving pre-test, post-test, and retention test was employed. In the experimental group, listening activities designed using LyricsTraining were implemented, while the control group followed the activities included in the 7th grade English textbook provided by the Ministry of National Education. The data were collected through a Listening Achievement Test and a Listening Anxiety Scale, and analyzed using Mixed Design ANOVA and ANCOVA. According to the results, there was no statistically significant difference between the post-test scores of the experimental and control groups. However, the results of the retention test revealed that the experimental group using LyricsTraining demonstrated more effective retention of listening skills compared to the control group. In addition, a statistically significant decrease was observed in the post-test listening anxiety scores of the experimental group. This indicates that LyricsTraining is an effective tool in reducing students' listening anxiety. The findings suggest that LyricsTraining enhances the long-term retention of listening skills and reduces anxiety levels among students. In this regard, it is recommended that LyricsTraining be used more frequently and at regular intervals to increase its effectiveness.

Keywords: Foreign language teaching, listening achievement, listening skills, listening anxiety, LyricsTraining.

Citation: Genç, U., & Acar Erdol, T. (2025). The Effect of LyricsTraining on Listening Achievement and Listening Anxiety in Foreign Language Education. Ankara University Journal of Faculty of Educational Sciences, 58(2), 819-864. <https://doi.org/10.30964/aujebfd.1651336>

¹This study is derived from the master thesis completed by the first author under the supervision of the second author.

²Teacher, Ministry of National Education, Türkiye, E-mail: uurgenc83@gmail.com, <https://orcid.org/0009-0002-9748-7830>, <https://ror.org/00jga9g46>

³Corresponding Author: Assoc. Prof. Dr., Ordu University, Faculty of Education, Department of Curriculum and Instruction, Ordu/Türkiye, E-mail: tubaacarerdol@gmail.com, <https://orcid.org/0000-0002-6954-4968>, <https://ror.org/04r0hn449>

Foreign language learning requires the functional integration of four fundamental language skills (listening, speaking, reading, and writing) along with three components (grammar, pronunciation, and vocabulary) (Brown, 2007). In theory, each skill and component is considered equally important and should be harmoniously integrated. However, in practice, listening skills are often neglected and remain in the background during foreign language instruction (Kahraman, 2015). This is primarily due to various factors that hinder the recognition of listening as equally important as other skills—such as perceptual difficulties during listening, speech rate, accent, vocabulary knowledge, verbal and non-verbal feedback, and learners' background knowledge (Hamouda, 2013).

Listening plays a pivotal role in the development of other language skills (Nunan, 1999), as individuals tend to speak the way they hear, and their overall language development depends on their ability to comprehend what they hear (Krashen, 1996). The development of listening skills in a foreign language is as crucial as in one's native language. However, in Türkiye, teachers often strictly adhere to textbooks and prioritize other types of activities included in them, resulting in insufficient focus on listening practices. In order to improve listening skills, a diversification of activities is required. There exists a notable discrepancy between the objectives and outcomes in foreign language education in Türkiye, particularly regarding the development of listening skills (Avcı & Hol, 2023). Kahraman (2015) examined the underlying reasons for weak listening skills in Türkiye and emphasized the shortcomings in the educational process and inadequacy of instructional materials. He highlighted that listening activities are often insufficient and not tailored to students' proficiency levels, which adversely affects their ability to construct meaning and develop language skills. Furthermore, he stated that teachers' dependency on textbooks limits the use of innovative methods and results in an approach that disregards individual learner differences.

One of the major problems faced by students in foreign language learning is listening anxiety—a psychological barrier that can impede the development of listening skills (Elkhafaifi, 2005; Horwitz et al., 1986). Factors such as students feeling shy or tense when speaking in class, avoiding answering teachers' questions due to lack of self-confidence, struggling to choose appropriate words and grammatical structures while speaking, and teachers delivering monotonous lessons all negatively affect language skill development (Chau, 2021). Listening anxiety is considered a critical variable that significantly influences listening skill development and academic performance. Research has shown a strong negative correlation between listening anxiety and listening proficiency (Elkhafaifi, 2005; Polat, 2019; Zhang, 2013). Students with high levels of listening anxiety tend to use fewer listening strategies, have lower comprehension levels, and consequently perform poorly overall. Conversely, students with lower listening anxiety have been observed to use metacognitive strategies more effectively (Vandergrift, 2007).

Listening anxiety occurs across all levels of education. Among middle school students, anxiety related to listening skills has become a significant focus of recent educational research. Numerous studies have been conducted on this issue in Türkiye as well (Ay & Akçay, 2024; Orakçı, 2018; Say Yağcıoğlu & Çalışkan, 2024; Yükselir & Özer, 2022). Bayraktar et al. (2022) noted that the increasing number of studies on listening skills at the middle school level highlights both the prevalence of related problems and the growing interest of researchers in this topic. Difficulties in understanding texts and keeping up with fast speech are among the main causes of listening anxiety in middle school students (Orakçı, 2018). This anxiety hinders students' active participation in class and negatively impacts the learning process (Melanlıoğlu, 2013). Arslan (2018) emphasized that the anxiety experienced by middle school students while learning a foreign language adversely affects their academic achievement.

This study focused on middle school students, primarily because the 10–14 age group represents a critical developmental period for language learning. At this stage, students experience rapid cognitive development, enhanced processing of linguistic input, and the shaping of attitudes toward foreign languages (Demirezen, 2003). The middle school period is a critical phase in which the foundation of language skills is laid and future learning processes are shaped (Demirezen, 2003). It is noted that students in this age group exhibit higher motivation toward language learning, and the skills acquired during this period directly affect subsequent learning (Birdsong, 2018; Ocak et al., 2022). Furthermore, middle school students show a strong interest in technology, making digitally supported learning activities highly effective during this phase (Birdsong, 2018). OECD (2021) reports also indicate that countries that prioritize language education in middle school curricula tend to have higher long-term academic success among students. Kemiksiz (2017) emphasized that the lack of effective listening instruction in middle school may hinder the development of listening skills in later years. Therefore, it is essential to identify effective methods to reduce listening anxiety and improve students' listening performance.

Krashen (1996) stated that encountering real-life examples during listening practice supports the language learning process. Tools like LyricsTraining not only provide learners with authentic language input but also create a motivating learning environment. Through song lyrics, learners can naturally develop their grammar knowledge, vocabulary, and pronunciation skills.

LyricsTraining, with its interactive structure, is expected to reduce students' fear of making mistakes and add an element of fun to the learning process. The platform can be used both for individual study and in group activities within the classroom, promoting autonomous learning as well as collaborative learning experiences. Its adaptability to different proficiency levels also makes it an effective tool for learners at various stages.

From a teacher's perspective, the flexibility offered by tools like LyricsTraining enables the design of customized content for language teaching. For example, teachers

can select topic-specific vocabulary to create targeted activities that meet students' needs. In addition, the platform provides a system that allows learners' performance to be measured, making the language learning process more structured and trackable for both students and teachers (Garcia, 2015). For this reason, LyricsTraining has been the subject of various studies.

When reviewing the literature, it is observed that there are a limited number of international studies investigating the impact of LyricsTraining on listening skills. Most of these studies have been conducted at the university level (Angelina, 2020; Darmawan & Prischilla, 2019; Gamboa Galarza, 2022; Manurung et al., 2022; Meilinda, 2020; Nguyen, 2023), and only two studies have been carried out at the high school level (Paras, 2024; Veloz Toapanda, 2023). However, there appears to be no study in the Turkish context examining the effects of LyricsTraining on listening performance and anxiety. English language instruction in Türkiye differs significantly from foreign contexts in terms of cultural dynamics, curriculum structure, and pedagogical practices. Therefore, investigating the effects of LyricsTraining on Turkish students is crucial for evaluating the tool's effectiveness within the local context. Moreover, such a study has the potential to provide teachers with concrete insights into the classroom application of technology-supported listening activities. Conducting an experimental study at the middle school level is particularly valuable, considering students' familiarity with digital tools and the prevalence of listening anxiety during foreign language learning. Such a study would fill an important gap in the literature.

The aim of this study is to determine the effects of using LyricsTraining in seventh-grade foreign language classes on students' listening performance and listening anxiety. To this end, the following research questions are addressed:

1. Is there a significant difference between the experimental and control groups in terms of the changes in Listening Achievement Test (LAT) mean scores from pre-test to post-test and retention test?
2. When the pre-test results of the Listening Anxiety Scale (LAS) are controlled, is there a significant difference between the experimental and control groups in their post-test scores?

Method

Research Model

This study employed a quasi-experimental design involving a pre-test (PT), post-test (PoT), and retention test (RT), with comparisons made between an experimental group (EG) and a control group (CG). This design is used to assess whether the observed differences between groups are the result of the experimental intervention (Büyüköztürk, 2018). Accordingly, both groups were administered the PT at the outset, but the intervention was implemented only with the EG. Following the intervention, both groups were administered the PoT, and a subsequent RT was also

conducted. While the EG received listening instruction using LyricsTraining, the CG followed the 2018 English Curriculum and the accompanying textbook published by the Turkish Ministry of National Education (MoNE). In addition, both groups completed the Listening Achievement Test (LAT) and the Listening Anxiety Scale (LAS) before and after the intervention to examine changes in their listening skills and anxiety levels. The symbolic representation of the quasi-experimental design used in the study is presented in Table 1.

Table 1

Research Design

Group	Pre-Test	Intervention	Post-Test	Retention Test
EG	O ₁ (LAT & LAS)	Listening practice with LyricsTraining	O ₃ (LAT & LAS)	O ₅ (LAT)
CG	O ₂ (LAT & LAS)	7th Grade MoNE English Textbook	O ₄ (LAT & LAS)	O ₆ (LAT)

Study Group

The research was conducted with five different study groups, all consisting of middle school students from a mid-sized city in Türkiye. The first group consisted of 84 eighth-grade students and was used to conduct item analysis of the LAT. The second group included 150 eighth-grade students and was used to perform a Confirmatory Factor Analysis (CFA) of the LAS in English. The third group consisted of seventh-grade students from a single middle school. The LAT was administered to a total of 210 students across seven classes. The mean scores obtained by each class were as follows: A = 6.58, B = 7.02, C = 7.69, D = 7.97, E = 8.40, F = 6.10, G = 8.70. Based on these averages, two classes with the most similar scores (Class C = 7.69 and Class D = 7.97) were randomly assigned to serve as the experimental group and control group, forming the fourth and fifth study groups, respectively. To determine whether there was a statistically significant difference in the pre-test scores between the EG and CG, an independent samples t-test was conducted. The results revealed no statistically significant difference between the groups [$t(63) = 0.33$; $p > 0.05$]. The demographic data of the students in the EG and CG are presented in Table 2.

Table 2

Demographic and Academic Averages of Students in the Experimental and Control Groups

Variable	EG (n = 33, %)	CG (n = 32, %)
Gender		
Male	17 (51.5)	17 (53.13)
Female	16 (48.5)	15 (46.87)
Age		
11	1 (3.3)	3 (9.3)
12	29 (87.8)	26 (81.4)
13	3 (9.9)	3 (9.3)
Academic		
Average	81.44	84.87
English		
Average	81.44	80.20

As shown in Table 2, the EG and CG are similar in terms of gender and age distribution. In both groups, the number of female students is slightly higher than that of male students, and the differences between the groups are minimal. Regarding academic achievement, the CG has an average score that is 2 points higher than the EG. However, in terms of English course averages, the difference is only 0.5 points, indicating that both groups have relatively similar levels of English proficiency.

Data Collection Tools

In this study, two separate measurement tools were used to evaluate students' listening skills and listening anxiety.

Listening Achievement Test (LAT)

The LAT was developed by the researchers in alignment with the learning outcomes for the first semester of seventh grade as specified in the 2018 English Language Curriculum (MoNE, 2018). The test consists of multiple-choice questions, each with four answer options. During the development of the items, seventh-grade supplementary resources available on the Educational Informatics Network (EBA) platform were utilized. Based on these materials, the items were converted into English audio files using the Naraket application. In the 2018 Curriculum, each unit at the seventh-grade level includes at least one listening outcome. Since the research was conducted during the first semester of the 2023–2024 academic year, five themes

from the curriculum were selected, and questions were developed in accordance with one outcome from each unit.

To ensure the validity and reliability of the LAT, feedback was obtained from eight academic experts: three in English language education, four in curriculum and instruction, and one in educational measurement and evaluation. Based on their feedback concerning content validity, scientific accuracy, student level appropriateness, and item-writing principles, four items were removed and five were revised. As a result, the draft version of the test consisted of 25 items. A pilot implementation of the test was carried out with 84 eighth-grade students. This grade level was selected for the pilot because these students had previously been taught the content included in the listening test. The results were analyzed using the Test Analysis Program (TAP), and the resulting test statistics are presented in Table 3.

Table 3

Item Statistics for the Listening Achievement Test

Item No	Difficulty Index	Discrimination Index	Item No	Difficulty Index	Discrimination Index
01	0.9	0.25	14	0.61	0.51
02	0.88	0.25	15	0.5	0.59
03	0.71	0.58	16	0.62	0.59
04	0.65	0.55	17	0.57	0.8
05	0.48	0.76	18	0.76	0.71
06	0.27	0.68	19	0.61	0.79
07	0.36	0.68	20	0.54	0.8
08	0.27	0.64	21	0.79	0.54
09	0.39	0.68	22	0.7	0.71
10	0.26	0.48	23	0.69	0.67
11	0.14	0.11	24	0.67	0.67
12	0.5	0.63	25	0.75	0.58
13	0.73	0.22			

As shown in Table 3, the item difficulty indices range from 0.14 to 0.90. The first item, with the highest difficulty index (0.90), was answered correctly by the vast majority of students and is therefore considered very easy. In general, a higher difficulty index indicates that an item is easier, while lower values reflect a more

difficult item (Turgut & Baykul, 2018). The 11th item, with a difficulty index of 0.14, was the most challenging for students.

Item discrimination indices ranged from 0.11 to 0.80, suggesting that the majority of the items had high discriminating power. Items 17, 19, and 20 had the highest discrimination indices (up to 0.80), effectively distinguishing between high-performing and low-performing students. According to standard criteria, an index of 0.40 or above indicates high item discrimination (Büyüköztürk, 2018). Item 11 had both low difficulty (0.14) and low discrimination (0.11) and was therefore considered detrimental to the reliability of the test.

Based on item analysis, items with difficulty indices between 0.30 and 0.70 are considered moderately difficult and generally acceptable. Items below 0.30 are interpreted as difficult, and those above 0.70 as easy (Hasançebi et al., 2020; Turgut & Baykul, 2018;). For discrimination, the interpretation is as follows: ≥ 0.40 = high discrimination, $0.30-0.39$ = moderate, $0.20-0.29$ = low, and <0.20 = very low. Items with discrimination below 0.20 are recommended for removal (Hasançebi et al., 2020). Accordingly, the following items were removed from the test:

- **Item 1:** Very easy (difficulty = 0.90) and low discrimination (0.25). Removed due to limited evaluative capacity.
- **Item 2:** Very easy (difficulty = 0.88), low discrimination (0.25). Removed due to weak differentiation.
- **Item 8:** Difficult item (difficulty = 0.27) but high discrimination (0.64). Removed due to deviation from optimal difficulty range.
- **Item 10:** Very difficult (difficulty = 0.26), adequate discrimination (0.48). Removed due to risk of distorting overall test balance.
- **Item 11:** Very difficult (difficulty = 0.14) and very low discrimination (0.11). Removed due to negative effect on reliability.
- **Item 13:** Easy (difficulty = 0.73), low discrimination (0.22). Removed due to poor discriminative ability.

In addition, to ensure an equal number of questions per learning outcome, four other items with close values (Items 16, 18, 21, and 25) were also removed. As a result, the final version of the LAT consisted of 15 items. The test specification table (Table 4) presents the thematic distribution, learning outcomes, and cognitive levels of the finalized test items.

Table 4*Test Specification Table for the Listening Achievement Test*

Theme	Outcome Code	Outcome Description	Cognitive Level	No. of Items	Item Numbers
Appearance and Personality	E7.1.L1	Students will be able to understand clear, standard speech about appearance and personality.	Understand	3	1–3
Sports	E7.2.L1	Students will be able to recognize frequency adverbs in simple spoken texts.	Remember	3	4–6
Biographies	E7.3.L1	Students will be able to identify specific information about past events and dates.	Remember	3	7–9
Wild Animals	E7.4.L2	Students will be able to identify the names of wild animals in simple spoken texts.	Remember	3	10–12
Television	E7.5.L1	Students will be able to understand simple spoken texts about daily routines and preferences.	Understand	3	13–15

As seen in Table 4, the LAT consists of five themes. Each theme corresponds to a specific learning outcome and cognitive level. Three items were prepared for each theme to ensure balanced coverage.

Listening Anxiety Scale (LAS)

To measure students' English listening anxiety, the English Listening Anxiety Scale developed by Orakçı (2018) was used. This is a single-factor Likert-type scale with 14 items, rated on a five-point scale (1 = Strongly Disagree, 5 = Strongly Agree).

The reliability coefficient of the scale was reported as 0.987. Possible total scores range from 14 to 70. The fit indices of the scale were found to be at acceptable levels ($\chi^2/df = 2.04$, RMSEA = 0.056, AGFI = 0.941, RMR = 0.059, CFI = 0.97). Sample items include: “*I cannot understand English listening texts when I hear them for the first time*” and “*I have difficulty listening to English mass media (radio, television, CDs, etc.)*” (Orakcı, 2018).

Ethical Committee Approval

The planning and implementation of this study were approved by the Ethics Committee for Social Sciences and Humanities Research of the Rectorate of Ordu University with the decision dated 01/03/2023 and numbered 2023-107.

Implementation

The experimental study was conducted during the first semester of the 2023–2024 academic year. Prior to the experimental process, the Listening Achievement Test (LAT) was administered to 210 students across seven classes at the school where the researcher was employed. The students' test scores were analyzed, and the two classes with the most similar mean scores were randomly assigned as the experimental group (EG, $n=33$) and control group (CG, $n=32$).

In the first week of September, both groups were administered the pre-tests: the LAT and the LAS. Following this, the experimental intervention phase began. The EG received listening instruction using *LyricsTraining* for one class hour per week, totaling three hours per month. The listening activities were designed with *LyricsTraining* and implemented interactively as well as through printed worksheets.

In addition to the planned activities, students in the EG were encouraged to complete extra exercises from the *Top Lyrics* section of the *LyricsTraining* website (featuring the most popular lyrics globally) or based on the foreign artists they personally followed. This was aimed at increasing students' exposure to listening input and thus enhancing their listening skills.

In the CG, classes followed the MoNE English Language Curriculum (2018) using a pre-established lesson plan. Listening activities from the official textbook were implemented sequentially via smartboard. Sample activities used during the intervention are presented in the Appendix, and a summary of monthly implementations is shown in Table 5.

Table 5*Instructional Practices for Experimental and Control Groups*

Month / Lesson Hours	Theme	Experimental Group Activities	Control Group Activities
September 3 hours	Appearance and Personality	Listening activities designed with Lyrics Training were implemented.	Standard listening activities aligned with the theme were conducted based on the MoNE Curriculum.
October 3 hours	Sports	Listening activities designed with Lyrics Training were implemented.	Standard listening activities aligned with the theme were conducted based on the MoNE Curriculum.
November 3 hours	Biographies	Listening activities designed with Lyrics Training were implemented.	Standard listening activities aligned with the theme were conducted based on the MoNE Curriculum.
December 3 hours	Wild Animals	Listening activities designed with Lyrics Training were implemented.	Standard listening activities aligned with the theme were conducted based on the MoNE Curriculum.
January 3 hours	Television	Listening activities designed with Lyrics Training were implemented.	Standard listening activities aligned with the theme were conducted based on the MoNE Curriculum.

Data Analysis

The data were analyzed using the JAMOVİ and SPSS software programs. The Listening Anxiety Scale (LAS), originally developed for ninth-grade students, was subjected to CFA to assess its applicability to middle school students in this study. Data were collected from 150 seventh-grade students for the CFA. No missing data or univariate outliers were detected. To identify multivariate outliers, Mahalanobis distances were calculated, and as a result, the data from one participant were excluded from the analysis. Reliability analysis for the scale yielded a Cronbach's Alpha coefficient of 0.80. The CFA fit indices were calculated as follows: CFI=0.89, TLI=0.86, SRMR=0.063, and RMSEA=0.067. These results indicate that the scale is suitable for use with middle school samples.

In the analysis of the data, skewness and kurtosis values were examined to determine whether the variables were normally distributed within subgroups. According to George and Mallery (2010), skewness and kurtosis coefficients between -2 and +2 indicate a normal distribution. Similarly, Kline (2023) notes that values below an absolute value of 2 for skewness and 7 for kurtosis are considered acceptable for assuming normality. Descriptive statistics for the groups are presented in Table 6.

Table 6

Descriptive Statistics for the Groups

	Group	N	$\bar{x}$	df	Skewness	Kurtosis
LAT Pre-Test	EG	33	7.97	3.84	0.021	-1.01
	CG	32	7.69	3.02	0.161	-0.884
LAT Post-Test	EG	33	10.9	4.38	-0.743	-0.821
	CG	32	10.3	3.25	-0.449	-0.625
LAT Retention	EG	33	11.4	3.98	-1.05	-0.0646
	CG	32	10.3	3.68	-0.174	-1.43
LAS Pre-Test	EG	33	54.2	5.22	-0.496	-0.273
	CG	32	49.9	6.71	-1.07	1.09
LAS Post-Test	EG	33	43.6	6.47	-1.02	2.65
	CG	32	48.6	7.53	-0.953	0.486

Upon examining Table 6, it is observed that the skewness values across subgroups range between -1.07 and 0.161, while the kurtosis values range from -1.43 to 2.65. These values indicate that the data in the subgroups were normally distributed. To determine whether there were significant group and time differences in students' LAT scores between the experimental and control groups, a Mixed-Design ANOVA was conducted. Stevens (2002) emphasized that such analyses are advantageous in identifying both main effects and interaction effects. Brown and Forsythe (1974) noted that this method is a powerful tool for evaluating differences between independent groups and within-subject changes over time. Field (2018) also highlighted the importance of including time as a factor in understanding the effects of independent variables. Therefore, the Mixed-Design ANOVA was chosen as an appropriate method to effectively evaluate both group differences and performance changes over time in listening achievement.

When the assumptions of this test were examined, Mauchly's Test of Sphericity was conducted to test the equality of variances of the repeated measures. The results indicated that the assumption of sphericity was violated ($W = 0.808, p < .05$), meaning that the variances were significantly different. In cases where sphericity is violated, Greenhouse-Geisser and Huynh-Feldt corrections are recommended. The Greenhouse-Geisser epsilon was calculated as 0.839 and the Huynh-Feldt epsilon as 0.859. When the epsilon value exceeds 0.75, the Huynh-Feldt correction is preferred (Demirkol, 2024; Field, 2018). Therefore, the Huynh-Feldt adjusted results were taken into account in the analysis.

To examine whether there was a statistically significant difference between the experimental and control groups in terms of LAS scores, Analysis of Covariance (ANCOVA) was performed. ANCOVA is considered a powerful statistical method in experimental research, as it evaluates the effect of an independent variable while statistically controlling for the effects of one or more covariates (Field, 2018). In this study, the pre-test scores of the LAS were used as the covariate. Using the pre-test as a covariate helps eliminate the influence of initial differences between the experimental and control groups, thus providing a more accurate analysis. This approach aids in identifying the true effects of the independent variable on the dependent variable and prevents misleading conclusions (Tabachnick & Fidell, 2019). Additionally, taking the pre-test as a covariate acknowledges the initial differences between groups and enables their statistical control. This leads to a more precise evaluation of the experimental intervention's impact, thereby enhancing the internal validity of the study.

When the assumptions of ANCOVA were examined, Levene's Test revealed that the variances were homogeneous between groups ($p > .05$). However, the assumption of homogeneity of regression slopes was violated ($p < .05$). Nevertheless, Levy's (1980) simulation study showed that when group sizes are equal or nearly equal, ANCOVA is robust against violations of both the regression slope homogeneity and normality assumptions. In the present study, the assumption of normality was met, and group sizes were nearly equal (33 and 32). Moreover, ANCOVA was maintained due to its superiority over nonparametric alternatives (Hamilton, 1976).

Validity and Reliability Procedures

Factors that may affect the internal validity of a study include the measurement instrument, selection bias, subject attrition, testing effects, experimental procedures, and researcher influence (Fraenkel et al., 2012). In this study, one of the researchers also served as the teacher of the EG. The CG was taught by a different teacher who had comparable teaching experience. Considering that this could potentially introduce bias into the results, several measures were taken to ensure objectivity. First, the assessment tools administered to both the experimental and control groups were standardized instruments whose validity and reliability had been previously tested. The implementation process was carried out under identical physical and temporal conditions for both groups. Students were explicitly informed that the tests would not

affect their grades, that the evaluation process would remain confidential, and that the results would not be associated with their names. To preserve anonymity, all student data were coded as S1, S2, S3, etc., and no identifying information was used. Data analysis was conducted by the researcher who was not involved in the experimental implementation, in order to reduce potential bias. As previously mentioned, the two groups were selected randomly from among the seven groups who had taken the LAT, based on their similar mean scores—thus minimizing selection bias. No participant attrition occurred during the study. Additionally, ANCOVA was used as a statistical technique to enhance both the reliability and internal validity of the experimental research (Field, 2018).

Results

Results Regarding Changes in Listening Achievement

Results related to the first sub-question “Is there a statistically significant difference between the experimental and control groups in the changes in LAT scores from pre-test to post-test and retention test?” are presented in Table 7.

Table 7

Descriptive Statistics for the Experimental and Control Groups on the Listening Achievement Test

Measurement	Group	N	$\bar{x}$	df
LAT Pre-test	EG	33	7.97	3.84
	CG	32	7.69	3.02
LAT Post-test	EG	33	10.9	4.38
	CG	32	10.3	3.25
LAT Retention Test	EG	33	11.4	3.98
	CG	32	10.3	3.68

As seen in Table 7, the EG’s mean score increased from $\bar{x} = 7.97$ in the pre-test to $\bar{x} = 10.9$ in the post-test and further to $\bar{x} = 11.4$ in the retention test. Similarly, the CG’s mean score increased from $\bar{x} = 7.69$ in the pre-test to $\bar{x} = 10.3$ in the post-test and remained the same in the retention test. These results indicate that both groups started from comparable levels in terms of listening achievement. Although the EG received instruction supported by LyricsTraining, the post-test scores did not show a significant improvement over the traditional method applied in the CG. However, the retention scores suggest that the intervention had a more lasting effect on the listening skills of the EG.

To determine whether the differences in pre-test, post-test, and retention scores between the experimental and control groups were statistically significant, a Mixed-Design ANOVA was conducted. The results are presented in Table 8.

Table 8

Mixed-Design ANOVA Results for Listening Achievement

Source of Variance	Sum of Squares	df	Mean Square	F	p	Partial η^2
Between-Subjects						
Group (EG/CG)	21.4	1	21.4	0.596	0.443	0.009
Error	2265.5	63	36.0			
Within-Subjects						
Measurement	372.68	1.72	216.86	65.88	.001	0.511
Group $\times$ Measurement	5.70	1.72	3.32	1.01	.358	0.016
Error	356.36	108.26	3.29			

$p < 0.05^*$

As seen in Table 8, the interaction between group (EG/CG) and measurement time (pre-test, post-test, retention test) was not statistically significant ($F_{(3.32, 3.29)} = 1.01$, $p > .05$, $\eta^2 = 0.016$). However, a statistically significant difference was found across measurement times regardless of group ($F_{(216.86, 3.29)} = 65.88$, $p < .05$, $\eta^2 = .511$), indicating that time had a significant effect on listening achievement. To determine between which measurements the differences occurred, Bonferroni post-hoc comparisons were conducted. The results are presented in Table 9.

Table 9*Pairwise Comparisons of Listening Achievement Scores*

Group	Time (A)	Time (B)	Mean Difference (B-A)	Std. Error	P
EG	Pre-test	Post-test	2.97	0.414	.001*
		Retention	3.45	0.486	.001*
	Post-test	Retention	0.48	0.326	1.000
CG	Pre-test	Post-test	2.66	0.420	.001*
		Retention	2.63	0.494	.001*
	Post-test	Retention	-0.03	0.331	1.000

$p < 0.05^*$

As shown in Table 9, there were statistically significant differences between the pre-test and post-test and between the pre-test and retention test in both groups ($p < .05$). However, no significant difference was found between the post-test and retention test scores ($p = 1.000$), indicating that the level of performance attained in the post-test was maintained in the retention test.

Results Related to Changes in Listening Anxiety

Results regarding the second sub-question — “Is there a statistically significant difference between the experimental and control groups in post-test scores of the LAS when pre-test scores are controlled?” — are presented in Table 10.

Table 10*Descriptive Statistics for the Listening Anxiety Scale*

	Group	N	$\bar{x}$	Sd
LAS Pre-test	EG	33	54.2	5.22
	CG	32	49.9	6.71
LAS Post-test	EG	33	43.6	6.47
	CG	32	48.6	7.53

As shown in Table 10, the EG initially had a higher mean anxiety level ($\bar{x} = 54.2$) compared to the CG ($\bar{x} = 49.9$). At the end of the intervention, both groups

showed a reduction in anxiety. The mean post-test anxiety score decreased to 43.6 in the EG, while the CG's mean anxiety score was 48.6. The decrease in anxiety was more pronounced in the EG. The mean anxiety scores for listening, before and after adjustment according to the covariate, are presented in Table 11.

Table 11

Mean Listening Anxiety Scores of Groups Before and After Adjustment for Covariate

Group	N	$\bar{X}$	$\bar{X}$ (adjusted)
EG	33	43.6	42.5
CG	32	48.6	49.8

When Table 11 is examined, it is observed that when the pre-test scores of the LAS are controlled, the post-test scores of the EG, who received instruction enriched with Lyrics Training ($\bar{X} = 42.5$), are lower than those of the CG ($\bar{X} = 49.8$). The ANCOVA results related to the adjusted post-test mean scores based on the pre-test scores are presented in Table 12.

Table 12

ANCOVA Results Regarding Adjusted Post-Test Mean Scores Based on Pre-Test Scores

Source	SS	df	MS	F	p	Partial η^2
Group	766	1	766.3	19.3	<.001*	0.186
Pre-test (Covariate)	631	1	630.9	15.9	<.001*	0.151
Error	2466	62	39.8			

$p < 0.05^*$

As indicated in Table 12, ANCOVA revealed a statistically significant difference between the experimental and control groups in their adjusted post-test scores on the LAS ($F_{(1, 62)} = 19.3$, $p < .001$, $p\eta^2 = 0.186$). The calculated effect size ($p\eta^2 > .14$) indicates that the experimental intervention had a strong effect on reducing listening anxiety.

Discussion, Conclusion and Suggestions

This study investigated the effect of the LyricsTraining application on seventh-grade students' English listening achievement and anxiety. Post-test results showed that the LyricsTraining application did not create a statistically significant difference in improving listening skills. Although care was taken to use the application in accordance with learning outcomes and to carefully select content during the in-class implementation process, this result can be explained by the insufficient duration of the application and intervention. In this research, it is understood that the experimental intervention, applied for one hour per week over one semester, was not sufficiently effective. The literature emphasizes that the effects of song-based activities on language skills become more pronounced in the long term (Engh, 2013). Similar studies suggest that a minimum of 12 weeks of application yields significant results (Veloz Toapanda, 2023), and the weekly application time should be at least 2-3 hours. Furthermore, Ainun and Tahir's (2024) research on podcast-based language teaching states that the effect of technological tools becomes more pronounced with long-term and regular use. In light of these findings, it can be concluded that longer (at least one academic year) and more intensive (2-3 lessons per week) applications are necessary to fully measure the effectiveness of LyricsTraining.

However, retention test results indicate that LyricsTraining more effectively supported the sustainability of students' listening skills compared to the listening activities applied in the control group. This result supports the positive effects of music-based learning on long-term memory (Medina, 2002) and its ability to make learning permanent by forming emotional connections (Dolean, 2016). The increase in scores for both groups from pre-test to post-test also shows that the listening activities applied in the control group (using a smart board) provided a certain contribution to improving students' listening skills. This result demonstrates that traditional methods can also be effective, but digitally enhanced instruction provides more lasting learning.

One of the striking results of this research is that LyricsTraining was effective in reducing students' listening anxiety. This finding is consistent with previous research showing that the LyricsTraining application is effective in reducing students' listening anxiety (Angelina, 2020; Nguyen, 2023). Chien and Huang (2020) emphasize that digital tools, especially those that are enjoyable and easily accessible, play a significant role in reducing students' anxiety levels. Furthermore, Dolean (2016) highlights the power of songs not only to facilitate language learning but also to create a more positive learning environment. The music-based and interactive content offered by LyricsTraining allowed students to experience something different from traditional exam-oriented listening practices, enabling them to develop a more relaxed and positive attitude towards language learning. This reduction in students' anxiety levels can be explained by the interaction of multiple factors. Firstly, language input presented through songs provides a learning environment that attracts students' attention and creates emotional engagement. Enjoyable content reduced students' fear

of making mistakes, offering a safer learning environment. Additionally, activities in LyricsTraining such as fill-in-the-blanks, listening-based selection, and responding to multiple-choice questions have a task-based structure. This structure ensured students' active participation in the process and helped them focus on a specific goal. In task-based learning environments, as learners' progress can be observed step by step, this process also increases students' belief in their own competence (Okcan, 2023). The ability of students to instantly see their performance, receive immediate feedback from the system, and have the chance to repeat also contributed to reducing anxiety. In this process, the individualized and adaptable nature of learning allowed students to progress at their own pace and according to their needs, unlike traditional classroom settings. In this regard, it can be said that the song-based, enjoyable, and interactive environment offered by LyricsTraining reduced students' anxiety levels and made the foreign language learning process more efficient.

Recommendations

1. In future research, it is recommended to conduct studies with increased application duration and frequency, lasting at least two semesters and applying for more than one hour per week, to more clearly determine the effect of the LyricsTraining application.
2. This study examined the LyricsTraining application. Future research could aim to compare the effectiveness of different digital tools (e.g., VoScreen, Listenwise) to determine which tools are more effective in students' language learning processes.
3. The research was conducted on seventh-grade students. Studies with students from different age groups or different language proficiency levels could help evaluate the impact of such tools on a wider student population.
4. Given that the LyricsTraining application is effective in reducing students' listening anxiety, the longer and more regular use of this method will contribute to reinforcing the positive results obtained. It is recommended that teachers regularly integrate digital tools like LyricsTraining into their weekly lesson plans. Using the application for at least two lesson hours per week will facilitate students' more frequent exposure to language input and enable progress in their listening skills.
5. The LyricsTraining method, which is effective in reducing anxiety, can be supported with more interactive and enjoyable activities such as group work, competitions, and virtual reality applications. For example, students can compete as a team in class to complete song lyrics, or the group with the most correct answers within a certain period could be rewarded. Additionally, motivation can be increased by using feedback systems such as digital badges or achievement certificates, where students can observe their individual progress. Such activities can encourage students' active participation while also making the learning process more efficient and

enjoyable. Furthermore, student age levels, interests, and language proficiency should be considered when selecting songs, and content should be aligned with learning outcomes.

6. It is recommended that curriculum development specialists consider adding digital modules aimed at improving listening skills to English language curricula. In this regard, preparing application guides for teachers and presenting sample lesson plans supported with activity suggestions will help practitioners use digital tools more effectively.



Yabancı Dil Eğitiminde LyricsTraining'in Dinleme Başarısına ve Dinleme Kaygısına Etkisi¹

MAKALE TÜRÜ	Başvuru Tarihi	Kabul Tarihi	Yayın Tarihi
Araştırma Makalesi	05.03.2025	30.06.2025	16.08.2025

Uğur Genç ²

Milli Eğitim Bakanlığı

Tuba Acar Erdol ³

Ordu Üniversitesi

Öz

Bu araştırmanın amacı, LyricsTraining'in ortaokul 7. sınıf öğrencilerinin İngilizce dinleme başarıları ve dinleme kaygıları üzerindeki etkisini incelemektir. Araştırma, Türkiye'nin orta büyüklükteki bir ilinde bulunan bir ortaokulda, 2023-2024 eğitim-öğretim yılının birinci döneminde gerçekleştirilmiştir. Çalışmada, öntest, sontest ve kalıcılık testlerini içeren yarı deneysel desen kullanılmıştır. Deney grubunda LyricsTraining ile hazırlanan dinleme etkinlikleri uygulanırken, kontrol grubunda Milli Eğitim Bakanlığı'nın 7. sınıf İngilizce ders kitabında belirtilen etkinlikler kullanılmıştır. Veri toplama aracı olarak Dinleme Başarı Testi ve Dinleme Kaygı Ölçeği kullanılmış, veriler Karışık Desen ANOVA ve ANCOVA ile analiz edilmiştir. Araştırma sonuçlarına göre, deney ve kontrol gruplarının sontest puanları arasında anlamlı bir fark bulunmamıştır. Ancak, kalıcılık testinde LyricsTraining kullanılan deney grubunun, kontrol grubuna göre dinleme becerisinin kalıcılığını daha etkili bir şekilde desteklediği ortaya çıkmıştır. Ayrıca, deney grubunun dinleme kaygısı sontest puanlarında istatistiksel olarak anlamlı bir azalma gözlenmiştir. Bu durum, LyricsTraining'in öğrencilerin dinleme kaygısını azaltmada etkili bir araç olduğunu göstermektedir. Elde edilen bulgular, LyricsTraining'in dinleme becerilerinin kalıcılığını artırdığını ve öğrencilerin kaygı düzeylerini düşürdüğünü ortaya koymaktadır. Bu bağlamda, LyricsTraining'in etkisini artırmak için uygulamanın daha sık ve düzenli aralıklarla kullanılması önerilmektedir.

Anahtar sözcükler: Yabancı dil öğretimi, dinleme başarıları, dinleme becerisi, dinleme kaygısı, LyricsTraining.

¹Bu çalışma, birinci yazarın ikinci yazar danışmanlığında hazırladığı yüksek lisans tezinden üretilmiştir.

²Öğretmen, Milli Eğitim Bakanlığı, Türkiye, E-posta: uurgenc83@gmail.com, <https://orcid.org/0009-0002-9748-7830>, <https://ror.org/00jga9g46>

³Sorumlu Yazar: Doc. Dr., Ordu Üniversitesi, Eğitim Fakültesi, Eğitim Programları ve Öğretim Anabilim Dalı, Ordu/Türkiye, E-posta: tubaacarerdol@gmail.com, <https://orcid.org/0000-0002-6954-4968>, <https://ror.org/04r0hn449>

Yabancı dil öğrenimi, dört temel dilsel beceri (dinleme, konuşma, okuma, yazma) ve üç bileşenin (dilbilgisi, telaffuz ve kelime bilgisi) işlevsel bir bütünlük içinde kullanılmasını gerektirmektedir (Brown, 2007). Teoride, her bir beceri ve bileşenin eşit derecede önemli olduğu ve birbirine uyum sağlaması gerektiği vurgulanmaktadır. Ancak, uygulamada, yabancı dil öğretimi sırasında dinleme becerisi genellikle diğerlerine göre ihmal edilmekte ve arka planda kalmaktadır (Kahraman, 2015). Çünkü dinleme esnasında yaşanan algılama problemleri, konuşma hızı, aksan, kelime bilgisi, sözel ve sözel olmayan geri bildirimler, kişinin konu hakkındaki ön bilgisi gibi birçok faktör dinleme becerisinin diğer beceriler kadar önemsenmesinin önüne geçmektedir (Hamouda, 2013).

Dinleme becerisi diğer dilsel becerilerin gelişimi için kilit bir öneme sahiptir (Nunan, 1999). Çünkü kişi duyduğu gibi konuşur, duyduklarını anlamlandırabildiği ölçüde diğer becerileri gelişmektedir (Krashen, 1996). Yabancı dilde dinleme becerisinin gelişimi de ana dildeki gibi çok önemlidir. Ancak, Türkiye’de öğretmenler genellikle ders kitaplarına sıkı sıkıya bağlı kalmakta ve bu kitaplarda yer alan diğer etkinliklere daha fazla ağırlık verdikleri için dinleme etkinlikleri yeterince yapılamamaktadır. Dinleme becerisinin geliştirilmesi için etkinliklerin çeşitlendirilmesi gerekmektedir. Türkiye’de yabancı dil öğretiminde özellikle dinleme becerisinin geliştirilmesi konusunda hedefler ve sonuçlar arasında tutarsızlık vardır (Avcı ve Hol, 2023). Kahraman (2015), Türkiye’de dinleme becerilerinin zayıf olmasının ardındaki nedenleri inceleyerek, eğitim sürecindeki eksikliklerin ve kullanılan materyallerin yetersizliğinin bu durumda rol oynadığını belirtmektedir. Çalışmasında, dinleme etkinliklerinin genellikle yetersiz ve öğrencilerin dil seviyelerine uygun olmayan materyallerle desteklendiğini vurgulamaktadır. Bu durum, öğrencilerin anlam çıkarma ve dil becerilerini geliştirme süreçlerini olumsuz etkilemektedir. Ayrıca, öğretmenlerin ders kitaplarına bağımlı kalarak yenilikçi yöntemleri sınırlandırdıkları ve öğrencilerin bireysel farklılıklarını dikkate almayan bir eğitim yaklaşımı benimsediklerini ifade etmektedir.

Yabancı dil öğreniminde öğrencilerin karşılaştığı önemli sorunlardan biri de dinleme kaygısıdır. Dinleme kaygısı, öğrencilerin dinleme becerilerini geliştirmelerini engelleyebilecek duygusal bir bariyerdir (Elkhafaifi, 2005; Horwitz vd., 1986). Öğrencilerin sınıf ortamında konuşurken utangaç olmaları ve gergin hissetmeleri, öğretmenlerin sorularına özgüven eksikliği nedeniyle cevap vermekten kaçınmaları, konuşurken uygun kelimeleri ve doğru dil bilgisi kurallarını seçmeye çalışarak hata yapmaları ve öğretmenlerin dersleri monoton bir şekilde işlemeleri, öğrencilerin dil becerilerinin gelişimini olumsuz etkilemektedir (Chau, 2021). Yabancı dilde dinleme kaygısı, dinleme becerilerinin geliştirilmesi ve akademik başarı düzeyi üzerinde önemli bir etkisi olan kritik bir değişken olarak değerlendirilmektedir. Araştırmalar, dinleme kaygısı ile dinleme becerisi arasında negatif yönlü güçlü bir ilişki olduğunu ortaya koymuştur (Elkhafaifi, 2005; Polat, 2019; Zhang, 2013). Yüksek düzeyde dinleme kaygısı yaşayan öğrencilerin dinleme stratejilerini kullanma eğilimlerinin azaldığı, dinlediklerini anlama düzeylerinin düştüğü ve bu durumun genel başarılarına olumsuz yansıdığı belirtilmektedir. Ayrıca,

dinleme kaygısı düşük olan öğrencilerin, üstbilişsel stratejileri daha etkin bir şekilde kullandıkları gözlemlenmiştir (Vandergrift, 2007).

Dinleme kaygısı eğitimin her kademesinde yaşanmaktadır. Ortaokul öğrencilerinin yabancı dil öğrenme sürecinde yaşadıkları kaygıların, özellikle dinleme becerisiyle ilgili boyutları, günümüz eğitim araştırmalarının önemli bir odağı haline gelmiştir. Türkiye'de de bu konuya yönelik yapılan çalışmaların sayısı oldukça fazladır (Ay ve Akçay, 2024; Orakçı, 2018; Say Yağcıoğlu ve Çalışkan, 2024; Yükselir ve Özer, 2022). Bayraktar vd. (2022), ortaokul düzeyinde dinleme becerileri üzerine yapılan çalışmaların artışının, bu alandaki sorunların yaygınlığını ve araştırmacıların bu konuya verdiği önemi gösterdiğini belirtmiştir. Ortaokul öğrencilerinin dinleme etkinliklerinde metinleri anlama güçlüğü ve hızlı konuşmalara ayak uyduramamaları, dinleme kaygısının başlıca nedenlerindendir (Orakçı, 2018). Bu kaygı, öğrencilerin derse etkin katılımını engelleyerek öğrenme süreçlerini olumsuz etkilemektedir (Melanlıoğlu, 2013). Arslan (2018), ortaokul öğrencilerinin yabancı dil öğrenirken yaşadığı kaygıların, öğrenme başarılarını olumsuz etkilediğini belirtmektedir.

Bu çalışmada ortaokul öğrencileri ile çalışılmıştır. Bunun temel nedeni, 10–14 yaş grubunun yabancı dil öğrenimi açısından kritik bir gelişim dönemi olmasıdır. Bu yaş aralığında öğrencilerin bilişsel gelişimleri hız kazanmakta, dilsel girdileri işleme becerileri artmakta ve yabancı dile yönelik tutumları biçimlenmektedir (Demirezen, 2003). Ortaokul dönemi, dil becerilerinin temellerinin atıldığı ve ileriki öğrenme süreçlerini şekillendiren kritik bir aşamadır (Demirezen, 2003). Bu dönemdeki öğrencilerin dil öğrenimine karşı daha yüksek motivasyon sergilediği ve bu dönemde kazanılan becerilerin sonraki öğrenme süreçlerini doğrudan etkilediği belirtilmektedir (Birdsong, 2018; Ocak vd., 2022). Ayrıca, ortaokul çağındaki öğrenciler teknolojiye karşı yüksek bir ilgi göstermekte olup, bu dönemde dijital araçlarla desteklenen öğrenme etkinliklerinden yüksek verim alınabileceği öngörülmektedir (Birdsong, 2018). OECD (2021) raporları da, ortaokul öğretim programlarında dil eğitimine ağırlık veren ülkelerde öğrencilerin uzun vadeli akademik başarılarının daha yüksek olduğunu ortaya koymaktadır. Kemiksiz (2017), ortaokulda etkili dinleme eğitimi verilmemesinin ilerleyen yıllarda dinleme becerilerinin gelişimini olumsuz etkileyebileceğini vurgulamaktadır. Bu nedenle, dinleme kaygısını azaltmak ve öğrencilerin dinleme becerilerini geliştirmek için etkili yöntemlerin belirlenmesi gerekmektedir.

Krashen (1996), öğrencilerin dinleme pratiği yaparken gerçek yaşamdan örneklerle karşılaşmalarının dil öğrenme sürecini desteklediğini ifade etmektedir. Lyrics Training gibi araçlar, öğrencilere hem dil girdisi sağlamak hem de onları motive eden bir öğrenme ortamı sunmaktadır. Ayrıca, öğrenenlerin şarkı sözleri aracılığıyla dil bilgisi kurallarını, kelime dağarcığını ve telaffuz becerilerini doğal bir şekilde geliştirmelerine olanak tanımaktadır.

Etkileşimli bir yapı sunan Lyrics Training'in, öğrencilerin dil öğrenirken hata yapma korkusunu azaltması ve öğrenme sürecine eğlenceli bir boyut katması

beklenmektedir. Öğrencilerin bireysel olarak çalışabilecekleri gibi sınıf ortamında grup etkinlikleri şeklinde de kullanılabilen bu platform, hem bağımsız öğrenmeyi teşvik etmekte hem de işbirlikçi öğrenme deneyimlerine katkıda bulunmaktadır. Ayrıca, öğrencilerin seviyelerine göre uyarlanabilir bir yapıya sahip olması, bu aracı her düzeyde öğrenen için etkili bir hale getirmektedir.

Öğretmenler açısından da LyricsTraining gibi araçların sunduğu esneklik, dil öğretiminde özelleştirilmiş içerikler sunmayı mümkün kılmaktadır. Örneğin, öğretmenler belirli bir konuya yönelik kelimeleri seçerek öğrencilerin ihtiyaçlarına yönelik etkinlikler tasarlayabilmektedir. Bunun yanı sıra, uygulamanın öğrencilerin performanslarını ölçebilecekleri bir sistem sunması, hem öğrenciler hem de öğretmenler için dil öğrenme sürecini daha yapılandırılmış ve izlenebilir bir hale getirmektedir (Garcia, 2015). Bu nedenle LyricsTraining uygulaması çeşitli araştırmalara da konu olmuştur.

Yurt dışındaki araştırmalara bakıldığında, LyricsTraining uygulamasının dinleme becerisi üzerindeki etkilerini inceleyen sınırlı sayıda çalışmanın olduğu görülmektedir. Bu araştırmaların çoğunun üniversite düzeyinde olduğu tespit edilmiş (Angelina, 2020; Darmawan ve Prischilla, 2019; Gamboa Galarza, 2022; Manurung vd., 2022; Meilinda, 2020; Nguyen, 2023), yalnızca iki araştırmanın lise seviyesinde gerçekleştirildiği görülmüştür (Paras, 2024; Veloz Toapanda, 2023). Ancak LyricsTraining aracının dinleme başarısı ve kaygısı üzerindeki etkilerini inceleyen bir çalışmaya Türkiye bağlamında rastlanmamıştır. Türkiye'deki İngilizce öğretimi bağlamı, kültürel dinamikler, öğretim programı yapısı ve pedagojik uygulamalar açısından yurt dışındaki örneklerden belirgin biçimde farklılık gösterebilmektedir. Bu nedenle, LyricsTraining aracının Türkiye'deki öğrenciler üzerindeki etkilerinin incelenmesi, bu dijital aracın etkililiğini yerel bağlamda değerlendirmek açısından önemlidir. Ayrıca, böyle bir çalışma öğretmenlere, teknoloji destekli dinleme etkinliklerinin sınıf içi uygulamalarına dair somut ipuçları sunma potansiyeli taşımaktadır. Özellikle ortaokul düzeyinde yürütülecek deneysel bir araştırma, hem bu yaş grubunun dijital araçlara olan yatkınlığı hem de yabancı dil öğrenme sürecinde yaşadığı dinleme kaygısına yönelik çözüm üretme kapasitesi bakımından alanyazındaki önemli bir boşluğu dolduracaktır.

Bu araştırmanın amacı ortaokul yedinci sınıf yabancı dil dersinde LyricsTraining kullanımının ortaokul öğrencilerinin dinleme başarılarına ve kaygılarına etkisini belirlemektir. Bu amaçla şu sorulara yanıt aranmıştır:

1. Dinleme Başarı Testi (DBT) puan ortalamalarının uygulama öncesi ile uygulama sonrası ve kalıcılık ölçümleri değişimlerinde deney ve kontrol grupları arasında anlamlı bir farklılık var mıdır?
2. Dinleme Kaygı Ölçeği (DKÖ) ön test sonuçları kontrol edildiğinde deney ve kontrol gruplarının son test puanları arasında anlamlı bir farklılık var mıdır?

Yöntem

Araştırma Modeli

Bu araştırmada, ön test (ÖT), son test (ST) ve kalıcılık testi (KT) içeren, deney grubu (DG) ve kontrol grubu (KG) arasında karşılaştırmalar yapılan bir yarı deneysel desen uygulanmıştır. Bu desen, gruplar arasında gözlemlenen farklılıkların, deneysel müdahalenin bir sonucu olup olmadığını değerlendirmek için kullanılmaktadır (Büyüköztürk, 2018). Bu bağlamda, başlangıçta her iki gruba da ÖT uygulanmış, yalnızca DG'ye müdahale edilmiştir. Müdahale yalnızca DG'ye yapıp, müdahalenin ardından her iki gruba da ST uygulanmış ve sonrasında KT yapılmıştır. DG'de Lyrics Training kullanılırken, KG'de 2018 İngilizce Öğretim Programı ve bu doğrultuda hazırlanan ders kitabı takip edilmiştir. Ayrıca, her iki gruba araştırma öncesi ve sonrasında DBT ve DKÖ uygulanarak dinleme becerileri ve kaygı seviyelerindeki değişim incelenmiştir. Araştırmada tercih edilen yarı deneysel desene ait sembolik gösterime Tablo 1'de yer verilmiştir.

Tablo 1

Araştırmanın Deseni

Grup	Ön Test	İşlem	Son test	Kalıcılık Testi
DG	O ₁ DBT ve DKÖ	Lyrics Training ile dinleme çalışmaları	O ₃ DBT ve DKÖ	O ₅ DBT
KG	O ₂ DBT ve DKÖ	MEB Yedinci Sınıf İngilizce Ders Kitabı	O ₄ DBT ve DKÖ	O ₆ DBT

Çalışma Grubu

Araştırma beş çalışma grubu üzerinde yürütülmüştür. Çalışma gruplarını Türkiye'nin orta büyüklükteki bir ilinde yer alan ortaokul öğrencileri oluşturmuştur. Birinci çalışma grubu, DBT'nin madde analizlerini yapmak amacıyla oluşturulmuştur. Bu grupta sekizinci sınıf düzeyinde 84 öğrenci yer almıştır. İkinci grup, İngilizce DKÖ'nün doğrulayıcı faktör analizini (DFA) yapmak amacıyla oluşturulmuş olup, bu grupta sekizinci sınıf düzeyinde 150 öğrenci yer almıştır.

Üçüncü çalışma grubunu, bir ortaokulun yedinci sınıf düzeyinde yer alan öğrenciler oluşturmuştur. Yedi şubede yer alan 210 öğrenciye DBT uygulanmıştır. Öğrencilerin yapılan testlerden aldıkları puan ortalamaları hesaplanmıştır (A= 6.58, B= 7.02, C=7.69, D=7.97, E= 8.4, F=6.1, G=8.7) ve en yakın puanları alan iki şube (C=7.69 ve D=7.97) DG ve KG olarak rastgele belirlenerek araştırmanın dördüncü ve beşinci çalışma grupları oluşturulmuştur. DG ve KG'nin ÖT puanları arasındaki

anlamlı bir farklılık olup olmadığını test etmek için bağımsız örneklem t testi yapılmıştır. Bağımsız örneklem t testi sonuçları da DG ile KG arasında istatistiksel olarak anlamlı bir fark bulunmadığını göstermektedir [$t_{(63)} = 0.33$; $p > 0.05$]. DG ve KG’de yer alan öğrencilere yönelik demografik bilgilere Tablo 2’de yer verilmiştir.

Tablo 2

Deney ve Kontrol Gruplarında Yer Alan Öğrencilerin Demografik ve Akademik Ortalama Bilgileri

Değişken	DG (n=33, %)	KG (n=32, %)
Cinsiyet		
Kadın	17 (51.5)	17 (53.13)
Erkek	16 (48.5)	15 (46.87)
Yaş		
11	1 (3.3)	3 (9.3)
12	29 (87.8)	26 (81.4)
13	3 (9.9)	3 (9.3)
Akademik		
Ortalama	81.44	84.87
İngilizce		
Ortalaması	81.44	80.20

Tablo 2’ de sunulan DG ve KG’nin demografik bilgileri karşılaştırıldığında, her iki grup arasında benzer cinsiyet ve yaş dağılımları gözlemlenmektedir. Cinsiyet oranlarına bakıldığında, her iki grupta da kadın öğrenci sayısı erkeklerden daha fazla olup, gruplar arasındaki farklar çok azdır. Akademik ortalamalar açısından, KG ile DG arasında 2 puanlık bir fark bulunmaktadır. Ancak, İngilizce dersi ortalamalarında gruplar arasında yalnızca 0.5 puanlık bir fark vardır, bu da her iki grubun İngilizce başarıları açısından oldukça benzer seviyelerde olduğunu göstermektedir.

Veri Toplama Araçları

Araştırma sürecinde, öğrencilerin dinleme becerilerini ve dinleme kaygısını değerlendirmek amacıyla iki ayrı ölçme aracı kullanılmıştır.

Dinleme Başarı Testi

Araştırmacılar tarafından geliştirilen DBT, 2018 İngilizce Dersi Öğretim Programı (İDÖP)’nda yer alan yedinci sınıf birinci dönem kazanımlarına uygun olarak hazırlanmıştır. Test, dört seçenek içeren çoktan seçmeli sorulardan oluşmaktadır. Test soruları hazırlanırken, Eğitim Bilişim Ağı (EBA) platformunda yer

alan yedinci sınıf İngilizce yardımcı kaynaklardan faydalanılmıştır. Bu kaynaklardan yararlanarak oluşturulan sorular, Naraket uygulaması kullanılarak İngilizce ses dosyalarına dönüştürülmüştür. 2018 İDÖP 'te yedinci sınıf seviyesindeki her ünite de bir dinleme becerisi kazanımı yer almaktadır. Araştırma süresi 2023-2024 eğitim öğretim yılının birinci dönemini kapsadığı için, öğretim programındaki beş tema temel alınarak, bu ünitelerle ilişkili beş kazanım doğrultusunda sorular hazırlanmıştır. DBT'nin geçerlik ve güvenirliğini sağlamaya yönelik, üç konu alanı uzmanı (İngilizce eğitimi), dört program geliştirme uzmanı ve bir ölçme ve değerlendirme uzmanı olmak üzere toplam sekiz akademisyenin görüşleri alınmıştır. Kapsam geçerliğine, bilimsel doğruluğa, öğrenci düzeyine ve madde yazma ilkelerine uygunluğuna yönelik alınan dönütler doğrultusunda taslak ölçme aracındaki dört soru çıkarılmış ve beş soru düzenlenmiştir. Bu sürecin sonunda taslak ölçme aracı 25 sorudan oluşmuştur. Oluşturulan testin 84 kişilik sekizinci sınıf öğrencisi ile pilot uygulaması yapılmıştır. Pilot uygulamada sekizinci sınıf öğrencilerinin seçilmesinin nedeni bir önceki eğitim yıllarında dinleme testine konu olan üniteleri işlemiş olmalarıdır. Test sonuçları TAP (Test Analysis Program) ile analiz edilmiştir. Yapılan analiz sonucunda elde edilen test istatistikleri Tablo 3'te verilmiştir.

Tablo 3

Dinleme Başarı Testine Yönelik Test İstatistikleri

Soru No	Madde Güçlük İndeksi	Madde Ayırcılık İndeksi	Soru No	Madde Güçlük İndeksi	Madde Ayırcılık İndeksi
01	0.9	0.25	14	0.61	0.51
02	0.88	0.25	15	0.5	0.59
03	0.71	0.58	16	0.62	0.59
04	0.65	0.55	17	0.57	0.8
05	0.48	0.76	18	0.76	0.71
06	0.27	0.68	19	0.61	0.79
07	0.36	0.68	20	0.54	0.8
08	0.27	0.64	21	0.79	0.54
09	0.39	0.68	22	0.7	0.71
10	0.26	0.48	23	0.69	0.67
11	0.14	0.11	24	0.67	0.67
12	0.5	0.63	25	0.75	0.58
13	0.73	0.22			

Tablo 3'te yer alan madde istatistikleri, testin her bir maddesi için güçlük ve ayıricılık indekslerini göstermektedir. Tabloya göre, madde güçlük indeksleri 0.14 ile 0.9 arasında değişmektedir. En yüksek güçlük indeksine sahip olan 1. madde (0.90), öğrencilerin büyük çoğunluğu tarafından doğru yanıtlanmış olup oldukça kolay bir madde olarak değerlendirilmektedir. Madde güçlük indeksi arttıkça, maddenin kolaylık düzeyinin yükseldiği kabul edilir; buna karşılık düşük değerler, maddenin zor olduğunu göstermektedir (Turgut & Baykul, 2018). En düşük güçlük indeksine sahip olan 11. madde (0.14) ise öğrenciler tarafından en az doğru yanıtlanan ve dolayısıyla testin en zor maddesi konumundadır.

Madde ayıricılık indeksleri ise 0.11 ile 0.80 arasında değişmektedir. Bu değerler, maddelerin büyük çoğunluğunun yüksek ayırt ediciliğe sahip olduğunu göstermektedir. Özellikle 17., 19. ve 20. maddelerin ayıricılık indeksleri 0.80'e kadar ulaşmakta ve bu maddelerin yüksek başarı düzeyindeki öğrenciler ile düşük başarı düzeyindeki öğrencileri etkili bir şekilde ayırt edebildiğini ortaya koymaktadır. Genel kabul gören ölçütlere göre, ayıricılık indeksinin 0.40 ve üzeri olması madde açısından yüksek ayırt edicilik düzeyine işaret etmektedir (Büyüköztürk, 2018). Bununla birlikte, 11. madde hem güçlük (0.14) hem de ayıricılık (0.11) açısından düşük değerlere sahip olup, testin güvenilirliğine olumsuz katkı sağlayabilecek bir madde olarak değerlendirilmektedir.

Tablo 3 incelendiğinde, testin maddelerinin farklı zorluk düzeylerine ve ayırt edicilik özelliklerine sahip olduğu görülmektedir. Madde güçlük indeksinin 0.30 ile 0.70 arasında olması, maddenin orta düzeyde zorluk içerdiğini ve genellikle uygun kabul edildiğini göstermektedir. 0.30'un altındaki değerler zor, 0.70'in üzerindeki değerler ise kolay madde olarak yorumlanmaktadır (Hasançebi vd., 2020; Turgut & Baykul, 2018). Madde ayıricılık indeksinde ise 0.40 ve üzeri değerler "yüksek ayırt edici", 0.30–0.39 arası "orta düzeyde ayırt edici", 0.20–0.29 arası "düşük düzeyde ayırt edici" ve 0.20'nin altındaki değerler "çok düşük ayırt ediciliğe" sahip olarak değerlendirilmektedir. Ayırt ediciliği 0.20'nin altında olan maddelerin testten çıkarılması önerilmektedir (Hasançebi vd., 2020).

Bu ölçütler doğrultusunda yapılan analizde aşağıdaki maddelerin testten çıkarılmasına karar verilmiştir:

1. madde: Güçlük indeksi 0.90 olup oldukça kolay bir madde olarak değerlendirilmiştir. Ayrıca ayıricılık indeksi de 0.25 ile düşük düzeydedir. Hem kolay olması hem de düşük ayırt ediciliği nedeniyle testten çıkarılmıştır.
2. madde: Güçlük indeksi 0.88 (çok kolay) ve ayıricılık indeksi 0.25 (düşük). Öğrencilerin büyük çoğunluğu tarafından doğru yanıtlanmış ancak grupları ayırma gücü zayıf olduğundan çıkarılmıştır.
8. madde: Güçlük indeksi 0.27 ile zor bir madde olup, ayıricılık indeksi 0.64 ile yüksek olsa da zorluk düzeyi ideal aralığın altında kaldığı için testin genel yapısına uygun bulunmamıştır.

10. madde: Güçlük indeksi 0.26 ve ayıricılık indeksi 0.48. Ayıricılığı yeterli olsa da çok zor olması nedeniyle öğrencilerin başarısını ölçmede dengesizlik yaratabileceği gerekçesiyle çıkarılmıştır.

11. madde: Güçlük indeksi 0.14 ve ayıricılık indeksi 0.11 ile hem çok zor hem de çok düşük ayırt edici bir madde olup, testin güvenilirliğini olumsuz etkileyeceği düşünülerek testten çıkarılmıştır.

13. madde: Güçlük indeksi 0.73 (çok kolay) ve ayıricılık indeksi 0.22 (düşük). Madde, çok kolay olmasının yanı sıra yeterli ayırt edicilik göstermediği için test dışı bırakılmıştır. Daha sonra her bir kazanım için soru sayılarında eşitliği sağlamak adına birbirine yakın değerli olan sorulardan 16., 18., 21., ve 25. maddelerin de testten çıkarılmasına karar verilmiştir. Böylece DBT'nin nihai hali 15 sorudan oluşmuştur. Nihai halde testteki maddelere ilişkin belirtke tablosu tablo 4'te sunulmuştur.

Tablo 4

Dinleme Başarı Testi Belirtke Tablosu

Tema	Kazanım Kodu	Kazanım Açıklaması	Bilişsel Düzey	Soru Sayısı	Soru Numaraları
Appearance and Personality (Görünüş ve Kişilik)	E7.1.L1	Öğrenciler görünüşler ve kişilikler hakkındaki net ve standart konuşmayı anlayabilecektir.	Anlama	3	1-3
Sports (Sporlar)	E7.2.L1	Öğrenciler basit sözlü metinlerdeki sıklık zarflarını tanıyabilecektir.	Hatırlama	3	4-6
Biographies (Biyografiler)	E7.3.L1	Öğrenciler geçmiş olaylar ve tarihlerle ilgili sözlü metinlerdeki belirli bilgileri tanıyabilecektir.	Hatırlama	3	7-9

(devam ediyor)

Tablo 4 (devam)

Wild Animals (Vahşi Hayvanlar)	E7.4.L2	Öğrenciler basit sözlü metinlerdeki yabancı hayvanların adlarını tespit edebilecektir.	Hatırlama	3	10–12
Television (Televizyon)	E7.5.L1	Öğrenciler günlük rutinler ve tercihler hakkındaki basit sözlü metinleri anlayabilecektir.	Anlama	3	13–15

Tablo 4 incelendiğinde DBT’nin beş temadan oluştuğu ve her tema ile ilgili kazanım ve bilişsel düzeylere yer verildiği görülmektedir. Belirtke tablosunda soru dağılımları incelendiğinde her bir tema için üç adet soru hazırlandığı görülmektedir.

Dinleme Kaygısı Ölçeği

Bu çalışmada, öğrencilerin İngilizce dinleme kaygılarını ölçmek amacıyla Orakcı (2018) tarafından geliştirilen İngilizce DKÖ kullanılmıştır. Likert tipi bir derecelendirme sistemi (1:Kesinlikle Katılmıyorum, 5:Kesinlikle Katılıyorum) ile değerlendirilen ölçek, tek faktörden oluşmakta ve toplam 14 madde yer almaktadır. Ölçekle elde edilen ölçümlerin güvenirlik katsayısı 0.987 olarak bulunmuştur. Ölçekten alınabilecek puanlar 14 ile 70 arasında değişmektedir. Ölçeğin uyum değerlerinin ($\chi^2/df=2.04$, RMSEA=0.056, AGFI=0.941, RMR=0.059, CFI=0.97) istenen düzeyde olduğu belirlenmiştir. Örnek maddeler arasında, “*İngilizce dinleme metinlerini ilk defa dinlediğimde anlayamıyorum*” ve “*Kitle iletişim araçlarını (radyo, televizyon, cd vb.) İngilizce olarak dinlemekte zorluk çekiyorum*” gibi ifadeler yer almaktadır (Orakcı, 2018).

Etik Kurul Kararı

Bu araştırmanın planlanması ve yürütülmesi, Ordu Üniversitesi Rektörlüğü Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu’nun 01/03/2023 tarih ve 2023-107 nolu kararı ile uygun bulunmuştur.

Uygulama

Deneyisel çalışma, 2023-2024 akademik yılının birinci döneminde gerçekleştirilmiştir. Deneyisel süreç öncesinde araştırmacının görev yaptığı okulda yedi şubede yer alan 210 öğrenciye DBT uygulanmıştır. Öğrencilerin yapılan testlerden aldıkları puanlar analiz edilmiştir ve puanları birbirine en yakın olan iki şube, DG (n=33) ve KG (n=32) olarak rastgele belirlenmiştir. Belirlenen gruplara

Eylül ayının ilk haftasında ÖT olarak DBT ve DKÖ uygulanmıştır. Daha sonra deneysel müdahale süreci başlamıştır. DG’de, haftada bir saat olmak üzere ayda toplamda üç ders saati Lyrics Training ile hazırlanan dinleme etkinlikleri yapılmıştır. Etkinlikler hem interaktif olarak hem de çalışma kâğıtları hazırlanarak yapılmıştır. Bunun yanı sıra hazırlanan etkinliklerden farklı olarak DG’de yer alan öğrencilerden Lyrics Trainig sitesinde yer alan Top Lyrics (Dünya genelinde en çok beğenilen şarkı sözleri) başlığı altında yer alan veya öğrencilerin bireysel olarak takip ettiği yabancı şarkıcılara ait etkinlikleri yapmaları istenmiştir. Burada öğrencilerin daha fazla etkinliğe maruz kalarak dinleme becerilerini geliştirmeleri amaçlanmıştır.

KG’de ise MEB İDÖP doğrultusunda hazırlanan plana göre ders işlenmiştir. Derslerde akıllı tahta kullanılarak kitapta yer alan dinleme etkinlikleri sırayla yapılmıştır. Uygulama sürecine yönelik etkinlik örneklerine ekler bölümünde yer verilmiştir ve aylık yapılan işlemlere yönelik bilgiler Tablo 5’te sunulmuştur

Tablo 5

Deney ve Kontrol Gruplarına Yönelik Uygulamalar

Ay / Ders Saati	Tema	Deney Grubu Uygulamaları	Kontrol Grubu Uygulamaları
Eylül 3 Saat	Appearance and Personality	Lyrics Training ile hazırlanan dinleme etkinlikleri yapılmıştır.	İDÖP’e uygun şekilde “Appearance and Personality” temasıyla ilişkili standart dinleme etkinlikleri yapılmıştır.
Ekim 3 Saat	Sports	Lyrics Training ile hazırlanan dinleme etkinlikleri yapılmıştır.	İDÖP’e uygun şekilde “Sports” temasıyla ilişkili standart dinleme etkinlikleri yapılmıştır.

(devam ediyor)

Tablo 5 (devam)

Kasım 3 Saat	Biographies	Lyrics Training ile hazırlanan dinleme etkinlikleri yapılmıştır.	İDÖP'e uygun şekilde "Biographies" temasıyla ilişkili standart dinleme etkinlikleri yapılmıştır.
Aralık 3 Saat	Wild Animals	Lyrics Training ile hazırlanan dinleme etkinlikleri yapılmıştır.	İDÖP'e uygun şekilde "Wild Animals" temasıyla ilişkili standart dinleme etkinlikleri yapılmıştır.
Ocak 3 Saat	Television	Lyrics Training ile hazırlanan dinleme etkinlikleri yapılmıştır.	İDÖP'e uygun şekilde "Television" temasıyla ilişkili standart dinleme etkinlikleri yapılmıştır.

Verilerin Analizi

Verilerin analizi JAMOVİ ve SPSS programlarında yapılmıştır. DKÖ, dokuzuncu sınıf öğrencilerine yönelik geliştirilmiş olup, bu çalışmada ortaokul seviyesindeki öğrencilere uygulanabilirliğini test etmek amacıyla DFA yapılmıştır. DFA yapmaya yönelik 150 yedinci sınıf öğrencisinden veri toplanmış, kayıp veri ve tek değişkenli uç değer olmadığı tespit edilmiştir. Çok değişkenli uç değerleri belirlemek amacıyla Mahalanobis uzaklığı hesaplanmış ve bunun sonucunda bir katılımcının verisi çıkarılmıştır. Ölçeğe yönelik yapılan güvenirlik analizinde Cronbach's Alpha değeri 0.80 bulunmuştur. DFA analizine ait uyum indekslerinin CFI=0.89; TLI=0.86; SRMR=0.063 ve RMSEA=0.067 bulunmuştur. Bu veriler sonucunda ölçeğin ortaokul örnekleminde kullanılabilir olduğu görülmüştür.

Verilerin analizinde, değişkenlerin alt gruplarda normal dağılım gösterip göstermediklerini kontrol etmek için çarpıklık ve basıklık değerleri incelenmiştir. George ve Mallery (2010)'ye göre çarpıklık ve basıklık katsayıları -2 ile +2 arasında ise dağılım normal kabul edilmektedir. Bununla birlikte Kline (2023) çarpıklık için mutlak 2 ve basıklık için mutlak 7'nin altındaki değerlerin normal dağılım kabul edilebileceğini belirtmektedir. Gruplara ait betimsel istatistik değerlerine Tablo 6' da yer verilmiştir.

Tablo 6*Gruplara Ait Betimsel İstatistikler*

	Grup	N	$\bar{x}$	Ss	Çarpıklık	Basıklık
DBT Ön Test	DG	33	7.97	3.84	0.0211	-1.01
	KG	32	7.69	3.02	0.161	-0.884
DBT Son Test	DG	33	10.9	4.38	-0.743	-0.821
	KG	32	10.3	3.25	-0.449	-0.625
DBT Kalıcılık	DG	33	11.4	3.98	-1.05	-0.0646
	KG	32	10.3	3.68	-0.174	-1.43
DKÖ Ön Test	DG	33	54.2	5.22	-0.496	-0.273
	KG	32	49.9	6.71	-1.07	1.09
DKÖ Son Test	DG	33	43.6	6.47	-1.02	2.65
	KG	32	48.6	7.53	-0.953	0.486

Tablo 6 incelendiğinde alt gruplardaki çarpıklık değerlerinin -1.07 ile 0.161 arasında, basıklık değerlerinin ise -1.43 ile 2.65 arasında değiştiği görülmektedir. Bu değerler sonucunda verilerin alt gruplarda normal dağıldığı söylenebilir. DG ve KG'nin DBT'ye yönelik grup ve zaman farklarını belirlemek amacıyla Karışık Desen ANOVA yapılmasına karar verilmiştir. Stevens (2002), bu tür analizlerin ana etkiler ve etkileşimleri belirlemede avantaj sağladığını vurgulamaktadır. Brown ve Forsythe (1974), bu yöntemin, bağımsız gruplar arasındaki farkları ve zaman içindeki değişimleri değerlendirmek için güçlü bir araç olduğunu belirtmektedir. Field (2018), zaman faktörünün bağımsız değişkenlerin etkilerini anlamada önemli bir rol oynadığını ifade etmektedir. Bu nedenle, dinleme başarı testinde hem grup farklarını hem de zamanla değişen başarıyı daha etkili bir şekilde değerlendirmek amacıyla Karışık Desen ANOVA kullanılmıştır.

Bu testin varsayımları kontrol edildiğinde, ölçüm farklarının varyanslarının eşitliği koşulu Mauchly Küresellik testi ile sınanmış ve varyanslar arasında anlamlı bir fark olduğu için (Mauchly's $W=0.808$; $p<.05$) küresellik koşulunun sağlanmadığı görülmüştür. Küresellik varsayımı karşılanmadığında Greenhouse-Geisser ve Huynh-Feldt kestirimleri kullanılabilir. Küresellik testi sonucunda Greenhouse-Geisser epsilon değeri 0.839 ve Huynh-Feldt epsilon değeri 0.859 olarak hesaplanmıştır Epsilon değerinin 0.75'ten büyük olduğu durumlarda Huynh-Feldt

kestiriminin kullanılması önerilmektedir (Demirkol, 2024; Field, 2018). Bu nedenle analizde Huynh-Feldt testi sonuçları dikkate alınmıştır.

DG ve KG’de kaygı ölçeğinin ölçüm sonuçları arasında anlamlı bir farklılık olup olmadığının incelenmesinde, ANCOVA yapılmasına karar verilmiştir. ANCOVA, deneysel araştırmalarda güçlü bir istatistiksel yöntem olarak kabul edilmektedir. Bu yöntem, bağımsız değişkenin etkisini değerlendirirken aynı zamanda kovaryans olarak tanımlanan bir veya daha fazla kontrol değişkeninin etkisini ortadan kaldırarak sonuçların daha güvenilir olmasını sağlar (Field, 2018). DG ve KG’nin DKÖ öntest sonuçları kovaryans olarak alınmıştır. Öntest sonuçlarının kovaryans olarak kullanılması, DG ve KG arasında başlangıçtaki farklılıkların etkisini ortadan kaldırarak daha doğru bir analiz yapılmasını sağlamaktadır. Bu yaklaşım, bağımlı değişken üzerindeki gerçek etkilerin belirlenmesine yardımcı olur ve yanıltıcı sonuçların önüne geçer (Tabachnick ve Fidell, 2019). Ayrıca, öntestin kovaryans olarak alınması, gruplar arasında başlangıç seviyelerinin eşit olmadığını kabul ederek bu farkların kontrol edilmesini mümkün kılar. Böylece, deneysel müdahalenin etkisi daha hassas bir şekilde değerlendirilebilir ve bu da çalışmanın iç geçerliğini artırır. ANCOVA’nın varsayımları incelendiğinde, Levene testi sonuçlarına göre grupların varyanslarının homojen olduğu görülmüştür ($p > .05$). Ancak regresyon eğimlerinin homojenliği varsayımının karşılanmadığı tespit edilmiştir ($p < .05$). Bununla birlikte Levy (1980)’nin yaptığı simülasyon çalışmasında grup büyüklükleri eşit olduğunda ANCOVA’nın regresyon eğimlerinin homojenliği ve normallik varsayımlarının her ikisinin ihlaline karşı dirençli olduğu görülmüştür. Bu çalışmada normallik karşılandığından ve grup büyüklükleri eşitliğe çok yakın olduğundan (32, 33) ve ANCOVA’nın parametrik olmayan karşıtlarına göre üstünlüğünden (Hamilton, 1976) dolayı ANCOVA’ya yönelik analize devam edilmiştir.

Geçerlik ve Güvenirlik Çalışmaları

Bir araştırmanın iç geçerliğini etkileyen faktörler olarak ölçme aracının etkisi, seçim yanlılığı, denek kaybı, test etkisi, deneysel işlem, uygulayıcı etkisi gibi faktörler sıralanmaktadır (Fraenkel vd., 2012). Bu araştırmanın uygulama sürecinde, araştırmacıdan biri aynı zamanda DG’nin öğretmeni olarak görev yapmıştır. KG’nin öğretmeni ise araştırmacı ile benzer öğretmenlik deneyimine sahip farklı bir öğretmendir. Bu durumun araştırma sonuçları üzerinde önyargı oluşturma riski taşıyabileceği dikkate alınarak, nesnelliği sağlamak amacıyla çeşitli önlemler alınmıştır. Öncelikle, deney ve kontrol gruplarına uygulanan ölçme araçları, geçerliği ve güvenirliliği önceden test edilmiş, standartlaştırılmış araçlardır. Uygulamalar her iki grupta da aynı fiziksel ve zamansal koşullarda gerçekleştirilmiş; öğrencilere testlerin notla ilişkilendirilmediği, değerlendirme sürecinin gizli tutulduğu ve sonuçların isimleriyle eşleştirilmeyeceği açıkça belirtilmiştir. Bu doğrultuda, tüm öğrenci verileri Ö1, Ö2, Ö3 şeklinde kodlanarak kimlik bilgileri gizlenmiştir. Bu araştırmada verilerin analizi deneysel uygulamayı yapmayan diğer araştırmacı tarafından gerçekleştirilerek yanlılık önlenmeye çalışılmıştır. Seçim yanlılığını önlemek amacıyla, yukarıda açıklandığı üzere, DBT uygulanan yedi grup arasından birbirine en yakın ortalamaya sahip iki grup rastgele DG ve KG olarak atanmıştır. Araştırmada denek kaybı

olmamıştır. Deneysel araştırmanın sonuçlarının güvenilirliğini ve iç geçerliğini arttıran bir yöntem olarak ANCOVA (Field, 2018) kullanılmıştır.

Bulgular

Dinleme Başarısındaki Değişime Yönelik Bulgular

“DBT puan ortalamalarının uygulama öncesi ile uygulama sonrası ve kalıcılık ölçümleri değişimlerinde DG ve KG arasında anlamlı bir farklılık var mıdır?” alt problemine ait bulgular Tablo 7’de verilmiştir.

Tablo 7

Deney ve Kontrol Gruplarının Dinleme Başarı Testine Yönelik Betimsel İstatistikleri

Dinleme Başarı Testi	Grup	N	$\bar{x}$	SS
DBT Ön Test	DG	33	7.97	3.84
	KG	32	7.69	3.02
DBT Son Test	DG	33	10.9	4.38
	KG	32	10.3	3.25
DBT Kalıcılık	DG	33	11.4	3.98
	KG	32	10.3	3.68

Tablo 7’de sunulan DBT’ye yönelik betimsel istatistikler incelendiğinde DG’nin öntest puan ortalaması $\bar{x}= 7.97$ iken sontest puan ortalaması $\bar{x}= 10.97$ ’ye kalıcılık testinde ise $\bar{x}= 11.4$ ’e yükseldiği görülmüştür. KG’nin öntest ortalaması $\bar{x}= 7.69$ iken sontestte bu ortalama $\bar{x}=10.3$ ’e yükselmiş ve kalıcılık testinde aynı değer korunmuştur. Bu sonuçlar birlikte değerlendirildiğinde, dinleme başarı öntest sonuçlarının DG ve KG arasında birbirine oldukça yakın olduğu anlaşılmaktadır. Ancak sontestler incelendiğinde, DG’ye verilen Lyricstraining ile desteklenmiş eğitimin, KG’ye uygulanan geleneksel yöntemlerle verilen eğitime kıyasla dinleme becerisinin gelişiminde belirgin bir farklılık yaratmadığı görülmektedir. Bununla birlikte, kalıcılık testi sonuçları, DG’ye verilen eğitimin dinleme becerisinin kalıcılığında daha etkili olduğunu ortaya koymaktadır. DG ve KG’nin öntest, sontest ve kalıcılık testleri arasında anlamlı bir farklılık olup olmadığını test etmek amacıyla Karışık Desen ANOVA testi yapılmıştır. Test sonuçları Tablo 8’de sunulmuştur.

Tablo 8*Karışık Desen ANOVA Sonuçları*

<i>Varyansın Kaynağı</i>	<i>Kareler Toplamı</i>	<i>sd</i>	<i>Kareler Ortalaması</i>	<i>F</i>	<i>p</i>	<i>Kısmi η^2</i>
Gruplar arası						
Grup (Deney/Kontrol)	21.4	1	21.4	0.596	0.443	0.009
Hata	2265.5	63	36.0			
Gruplar içi						
Ölçüm (öntest – sontest - kalıcılık)	372.68	1.72	216.86	65.88	.001	0.511
Grup*Ölçüm	5.70	1.72	3.32	1.01	.358	0.016
Hata	356.36	108.26	3.29			

 $p < 0.05^*$

Tablo 8 incelendiğinde grup (DG ve KG) ve ölçüm zamanlarının (öntest, sontest, kalıcılık testi) etkileşimlerinin ortak etkilerinin istatistiksel olarak anlamlı olmadığı gözlenmiştir ($F_{(3.32, 3.29)} = 1.01$, $p > .05$, $\eta^2 = 0.016$). Bununla birlikte ölçüm zamanları arasında anlamlı bir farklılık olduğu görülmektedir ($F_{(216.86, 3.29)} = 65.88$, $p < .05$, $\eta^2 = 0.511$). Söz konusu farklılığın hangi gruplar arasında olduğunu belirlemek için post-hoc analizi testlerinden bonferroni kullanılmıştır. DBT'ye ilişkin ikili karşılaştırmalar Tablo 9'da sunulmuştur.

Tablo 9*Dinleme Başarı Testine İlişkin İkili Karşılaştırmalar*

Grup	Ölçüm(A)	Ölçüm (B)	Ort. Farkı (B-A)	Std. Hata	p
Deney	Öntest	Sontest	2.9697	0.414	.001
		Kalıcılık	3.4545	0.486	.001
	Sontest	Kalıcılık	0.4848	0.326	1.000
Kontrol	Öntest	Sontest	2.6563	0.420	.001
		Kalıcılık	2.6250	0.494	.001
	Sontest	Kalıcılık	-0.0313	0.331	1.000

 $p < 0.05^*$

Tablo 9 incelendiğinde ikili karşılaştırma sonuçlarına göre DG'nin öntest ile sontest ve öntest ile kalıcılık testi sonuçları arasında anlamlı bir farklılık vardır ($p<.05$). Ancak, sontest ile kalıcılık testi arasında anlamlı bir fark tespit edilmemiştir ($p=1.000$). Bu durumda, sontestte elde edilen başarının kalıcılık testi ile sürdürüldüğü söylenebilir. KG'de öntest ile sontest ($p<.001$) ve öntest ile kalıcılık testi ($p<.001$) arasında anlamlı fark bulunmaktadır. Ancak, sontest ile kalıcılık testi arasında anlamlı bir fark bulunmamıştır ($p=1.000$). Bu durum, KG'de elde edilen başarının sontest ve kalıcılık testi arasında değişmediğini göstermektedir.

Dinleme Kaygısındaki Değişime Yönelik Bulgular

“Dinleme Kaygı Ölçeği öntest sonuçları kontrol edildiğinde DG ve KG'nin sontest puanları arasında anlamlı bir farklılık var mıdır?” alt problemine ait bulgular Tablo 10'da sunulmuştur.

Tablo 10

Deney ve Kontrol Gruplarının Dinleme Kaygı Ölçeğine Yönelik Betimsel İstatistikleri

	Grup	N	$\bar{x}$	Ss
DKÖ Öntest	DG	33	54.2	5.22
	KG	32	49.9	6.71
DKÖ Sontest	DG	33	43.6	6.47
	KG	32	48.6	7.53

Tablo 10 incelendiğinde, öntest sonuçlarına göre, DG'nin ortalama kaygı düzeyi ($\bar{X}=54.2$) KG'ye ($\bar{X}=49.9$) kıyasla daha yüksek iken, her iki grup da sontestte kaygı düzeylerinde bir azalma göstermiştir. DG'nin sontestteki ortalama kaygı düzeyi 43.6'ya düşerken, KG'nin kaygı düzeyi ise 48.6 olarak kaydedilmiştir. DG'de kaygı düzeyinde belirgin bir azalma görülürken, KG'de ise azalma daha sınırlı kalmıştır. Grupların kontrol değişkenine göre düzeltme öncesi ve sonrası dinlemeye yönelik kaygı puan ortalamaları Tablo 11'de sunulmuştur.

Tablo 11

Grupların Kontrol Değişkenine Göre Düzeltme Öncesi ve Sonrası Dinlemeye Yönelik Kaygı Puan Ortalamaları

Grup	N	$\bar{X}$	$\bar{X}$ (düzeltilmiş)
DG	33	43.6	42.5
KG	32	48.6	49.8

Tablo 11 incelendiğinde, öğrencilerin DKÖ öntest puanları kontrol altına alındığında, Lyricstraining ile zenginleştirilmiş öğrenim gören DG öğrencilerinin DKÖ sontest puanları ($\bar{X}=42.5$), KG'deki öğrencilerin puanlarından ($\bar{X}=49.8$) daha düşük çıkmıştır. Öntest puanlarına göre düzeltilmiş sontest puan ortalamalarına ilişkin ANCOVA sonuçları ise Tablo 12'de sunulmuştur.

Tablo 12

Ön Test Puanlarına Göre Düzeltilmiş Son Test Puan Ortalamalarına İlişkin ANCOVA Sonuçları

Kaynak	KT	sd	KO	F	P	Kısmi η^2
Grup	766	1	766.3	19.3	<.001	0.186
Ön Test (ortak değişken)	631	1	630.9	15.9	<.001	0.151
Hata	2466	62	39.8			

$p<0.05^*$

Tablo 12 incelendiğinde, ANCOVA sonucuna göre DG ve KG'nin düzeltilmiş DKÖ sontest puanları arasında istatistiksel olarak anlamlı bir fark bulunmaktadır ($F_{(1,62)}=19.3$; $p<0.05$, $\eta^2=0.186$). Hesaplanan etki büyüklüğü değeri ($\eta^2>.14$), deneysel işlemin dinleme kaygısını azaltmaya yönelik güçlü bir etkiye sahip olduğunu göstermektedir.

Tartışma, Sonuç ve Öneriler

Bu çalışmada, LyricsTraining uygulamasının yedinci sınıf öğrencilerinin İngilizce dinleme başarıları ve kaygıları üzerindeki etkisi incelenmiştir. ST sonuçları, LyricsTraining uygulamasının dinleme becerilerini artırmada istatistiksel olarak anlamlı bir fark yaratmadığını göstermiştir. Sınıf içi uygulama sürecinde uygulamanın kazanımlara uygun şekilde kullanılmasına ve içeriklerin özenle seçilmesine dikkat edilmiş olmasına rağmen ulaşılan bu sonuç, uygulama ve müdahale süresinin yetersiz kalmasıyla açıklanabilir. Bu araştırmada, bir dönem boyunca haftada bir saat

uygulanan deneysel müdahalenin yeterince etkili olmadığı anlaşılmaktadır. Alan yazında, şarkı temelli etkinliklerin dil becerileri üzerindeki etkilerinin uzun vadede daha belirgin hale geldiği vurgulanmaktadır (Engh, 2013). Benzer çalışmalarda en az 12 haftalık uygulamaların anlamlı sonuçlar verdiği (Veloz Toapanda, 2023), haftalık uygulama süresinin ise en az 2-3 saat olması gerektiği önerilmektedir. Ayrıca, Ainun ve Tahir'in (2024) podcast tabanlı dil öğretimi üzerine yaptığı araştırmada, teknolojik araçların etkisinin uzun süreli ve düzenli kullanımda daha belirgin hale geldiği ifade edilmektedir. Bu bulgular ışığında, LyricsTraining'in etkililiğini tam olarak ölçebilmek için daha uzun süreli (en az bir akademik yıl) ve yoğun (haftada 2-3 ders) uygulamaların gerekli olduğu sonucuna varılabilir.

Bununla birlikte, KT sonuçları, LyricsTraining'in KG'de uygulanan dinleme etkinliklerine kıyasla öğrencilerin dinleme becerilerinin sürdürülebilirliğini daha etkili bir şekilde desteklediğini ortaya koymaktadır. Bu bulgu müzik temelli öğrenmenin uzun vadeli bellek üzerindeki olumlu etkilerini (Medina, 2002) ve duygusal bağlantılar kurarak öğrenmeyi kalıcı hale getirdiğini (Dolean, 2016) desteklemektedir. Her iki grubun da ÖT'den ST'ye kadar puanlarını artırması, KG'de uygulanan (akıllı tahta kullanılarak) dinleme etkinliklerinin de öğrencilerin dinleme becerilerini geliştirmede belirli bir katkı sağladığını göstermektedir. Bu sonuç, geleneksel yöntemlerin de etkili olabileceğini ancak dijital araçlarla zenginleştirilmiş öğretimin daha kalıcı öğrenme sağladığını ortaya koymaktadır.

Bu araştırmanın dikkat çeken sonuçlarından biri, LyricsTraining'in öğrencilerin dinleme kaygısını azaltmada etkili olmasıdır. Bu bulgu, LyricsTraining uygulamasının öğrencilerin dinleme kaygısını azaltmada etkili olduğunu gösteren önceki araştırmalarla da örtüşmektedir (Angelina, 2020; Nguyen, 2023). Chien ve Huang (2020) dijital araçların, özellikle eğlenceli ve kolay erişilebilir olanların, öğrencilerin kaygı düzeylerini azaltmada önemli bir rol oynadığını vurgulamaktadır. Ayrıca, Dolean (2016), şarkıların yalnızca dil öğrenimini değil, aynı zamanda daha olumlu bir öğrenme ortamı yaratma gücünü de vurgulamaktadır. LyricsTraining'in sunduğu içeriklerin müzik temelli ve etkileşimli olması, öğrencilerin geleneksel sınav odaklı dinleme uygulamalarından farklı bir deneyim yaşamalarına imkân tanımış, böylece dil öğrenimine karşı daha rahat ve olumlu bir tutum geliştirmelerini sağlamıştır. Öğrencilerin kaygı düzeyinde yaşanan bu azalma, birden fazla faktörün etkileşimiyle açıklanabilir. Öncelikle, şarkılar aracılığıyla sunulan dil girdisi, öğrencilerin dikkatini çeken ve duygusal etkileşim yaratan bir öğrenme ortamı sağlamaktadır. Eğlenceli içerikler, öğrencilerin hata yapma korkusunu azaltarak daha güvenli bir öğrenme ortamı sunmuştur. Ayrıca, LyricsTraining'de yer alan boşluk doldurma, dinleyerek seçim yapma ve çoktan seçmeli sorulara yanıt verme gibi etkinlikler, görev temelli bir yapıya sahiptir. Bu yapı, öğrencilerin hem sürece aktif katılımını sağlamış hem de belirli bir hedef doğrultusunda odaklanmalarına yardımcı olmuştur. Görev temelli öğrenme ortamlarında, öğrenenin başarısı adım adım gözlemlenebilir hale geldiğinden, bu süreç öğrencinin kendi yeterliliğine olan inancını da artırmaktadır (Okcan, 2023). Öğrencilerin kendi performanslarını anlık olarak görebilmeleri, sistem tarafından anında geri bildirim almaları ve tekrar etme şansına

sahip olmaları da kaygıyı azaltan etkenler arasında yer almaktadır. Bu süreçte öğrenmenin bireyselleştirilmiş ve öğrenciye göre uyarlanabilir olması, geleneksel sınıf ortamlarının aksine öğrencinin hızına ve ihtiyacına göre ilerleme imkânı tanımıştır. Bu doğrultuda, LyricsTraining'in sunduğu şarkı temelli, eğlenceli ve etkileşimli ortamın, öğrencilerin kaygı düzeylerini azaltarak yabancı dil öğrenme sürecini daha verimli hâle getirdiği söylenebilir.

Öneriler

1. Gelecekteki araştırmalarda, LyricsTraining uygulamasının etkisinin daha net bir şekilde belirlenebilmesi için, uygulama süresi ve sıklığının artırıldığı, en az iki dönem süren ve haftada bir saatten fazla uygulamanın yapıldığı çalışmalar yapılması önerilmektedir.

2. Bu çalışmada LyricsTraining uygulaması incelenmiştir. Gelecek araştırmalar, farklı dijital araçların (örneğin, VoScreen, Listenwise) etkinliklerini karşılaştırarak hangi araçların öğrencilerin dil öğrenme süreçlerinde daha etkili olduğunu belirlemeyi hedefleyebilir.

3. Araştırma yedinci sınıf öğrencileri üzerinde gerçekleştirilmiştir. Farklı yaş gruplarındaki veya farklı dil seviyelerindeki öğrencilerle yapılacak çalışmalar, bu tür araçların geniş bir öğrenci kitlesindeki etkisini değerlendirmeye yardımcı olabilir.

4. LyricsTraining uygulamasının öğrencilerin dinleme kaygısını azaltmada etkili olduğu göz önüne alındığında, bu yöntemin daha uzun süreli ve düzenli olarak kullanılması, elde edilen olumlu sonuçların pekiştirilmesine katkı sağlayacaktır. Öğretmenlerin LyricsTraining gibi dijital araçları haftalık ders planlarına düzenli biçimde entegre etmeleri önerilmektedir. Uygulamanın haftada en az iki ders saati kullanılması, öğrencilerin hem dil girdisine daha sık maruz kalmalarını hem de dinleme becerilerinde ilerleme sağlamalarını kolaylaştıracaktır.

5. Kaygıyı azaltmada etkili olan LyricsTraining yöntemi, grup çalışmaları, yarışmalar ve sanal gerçeklik uygulamaları gibi daha fazla interaktif ve eğlenceli etkinliklerle desteklenebilir. Örneğin, öğrenciler sınıf içinde takım halinde yarışarak şarkı sözlerini tamamlayabilir veya belirli bir süre içinde en çok doğru cevabı veren grup ödüllendirilebilir. Ayrıca, öğrencilerin bireysel ilerlemelerini gözlemleyebildikleri dijital rozetler veya başarı sertifikaları gibi geri bildirim sistemleri kullanılarak, motivasyon artırılabilir. Bu tür etkinlikler, öğrencilerin aktif katılımını teşvik ederken, aynı zamanda öğrenme sürecini daha verimli ve eğlenceli hale getirebilir. Ayrıca, şarkı seçiminde öğrencilerin yaş düzeyleri, ilgi alanları ve dil seviyeleri dikkate alınmalı, içerikler kazanımlarla uyumlu şekilde belirlenmelidir.

6. Program geliştirme uzmanlarının, İngilizce öğretim programlarına dinleme becerilerini geliştirmeye yönelik dijital modüller eklemeleri önerilebilir. Bu doğrultuda, öğretmenlere yönelik uygulama kılavuzlarının hazırlanması ve etkinlik önerileriyle desteklenmiş örnek ders planlarının sunulması, uygulayıcıların dijital araçları daha etkili biçimde kullanmalarına yardımcı olacaktır.

References

- Ainun, A., & Tahir, M. (2024). The use of podcasts in improving students listening skills at university level. *International Journal of Business English and Communication*, 2(2), 65-68. <https://doi.org/10.26858/ijobec.v2i2.1919>
- Angelina, P. (2020). The strengths and drawbacks of lyricstraining implementation in basic listening course. *ELTR Journal*, 4(2), 117-124. <https://doi.org/10.37147/eltr.v4i2.69>
- Arslan, A. (2018). Ortaokul öğrencilerinin konuşma kaygıları ve akademik özyeterlik inançlarının çeşitli değişkenler açısından incelenmesi [Investigation of secondary school students' reading anxiety and academic self-efficacy beliefs in terms of various variables]. *International e-Journal of Educational Studies*, 2(3), 26-43. <https://doi.org/10.30900/kafkasegt.338051>
- Avcı, A., & Hol, D. (2023). DIGIFLYERS: A digital storytelling program to improve the listening performance of young learners. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 66, 446-474. <https://doi.org/10.21764/maeuefd.1114660>
- Ay, D., & Akçay, A. O. (2024). The relationship between 7th grade tudents' autonomy perceptions and foreign language anxiety in foreign language courses. *Türk Akademik Yayınlar Dergisi*, 8(1), 70-97. <https://dergipark.org.tr/en/pub/tayjournal/issue/83899/1182399>
- Bayraktar Ozgur, N., & Dogan Kahtalı, B. (2022). Examining secondary school students' listening anxiety in terms of different variables and determining the causes of anxiety. *International Online Journal of Educational Sciences*, 14(4), 1006-102. <https://iojes.net/index.jsp?mod=tammetin&makaleadi=&makaleurl=c2da78b5-1f2e-45cd-ad5a-0bd7c7953064.pdf&key=62329>
- Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology*, 9(81), 1-17. <https://doi.org/10.3389/fpsyg.2018.00081>
- Brown, H. D. (2007). *Principles of language learning and teaching*. Pearson Longman.
- Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346), 364-367. <https://doi.org/10.1080/01621459.1974.10482955>
- Büyüköztürk, Ş. (2018). *Sosyal bilimler için veri analizi el kitabı [Handbook of data analysis for the social sciences]*. Pegem Akademi.
- Chau, K. G. (2021). The effect of ICT on learners' speaking skills development. *International Journal of TESOL & Education*, 1(1), 22-29. <https://ijte.org/index.php/journal/article/view/4>

- Chien, C. C., Huang, Y., & Huang, P. (2020). YouTube videos on EFL college students' listening comprehension. *English Language Teaching*, 13(6), 96-103. <https://doi.org/10.5539/elt.v13n6p96>
- Darmawan, Y. Y., & Prischilla, D. (2019). The use of Lyricstraining.com on listening class of first semester students of English education study program teacher training and education faculty, Universitas Bandar Lampung. *Beyond Linguistika*, 1(2), 70–71. <https://doi.org/10.36448/bl.v1i2.1347>
- Demirezen, M. (2003). Yabancı dil ve anadil öğreniminde kritik dönemler [Critical periods in first and second language acquisition]. *TÖMER Dil Dergisi*, 118, 5-15. <https://www.arapcadeposu.com/wp-content/uploads/2017/yabancidil-anadil-ogreniminde-kritik-donemler.pdf>
- Demirkol, S. (2024). Tekrarlı ölçümler için tek yönlü varyans analizi ve Friedman testi [One-way repeated measures ANOVA and the Friedman Test], In A. Ş. Avşar (Ed.). *Jamovi ile Kolay İstatistik [A Practical Guide to Statistics with Jamovi]* (pp.107-127). Nobel Yayıncılık.
- Dolean, D. D. (2016). The effects of teaching songs during foreign language classes on students' foreign language anxiety. *Language Teaching Research*, 20(5), 638-653. <https://doi.org/10.1177/1362168815606151>
- Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal*, 89(2), 206-220. <https://doi.org/10.1111/j.1540-4781.2005.00275.x>
- Engl, D. (2013). Why use music in English language learning? A survey of the literature. *English Language Teaching*, 6(2), 113-127. <http://dx.doi.org/10.5539/elt.v6n2p113>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. Sage
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. McGraw-hill.
- Gamboa Galarza, D. C. (2022). *Lyrics training and listening skill*, [Bachelor's thesis, Universidad Técnica]. <https://repositorio.uta.edu.ec/items/895a315f-111d-4079-a242-8ca4147ea804>
- García, C. A. B. (2015). The use of LyricsTraining website to improve listening comprehension. *Mextesol Journal*, 39(3), 1-13. <https://www.mextesol.penamiller.com/static/public/files/b65ed0a845f0fa50e58726a3c858ad62.pdf>
- George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference*. Pearson

- Hamilton, B. L. (1976). A Monte Carlo test of the robustness of parametric and nonparametric analysis of covariance against unequal regression slopes. *Journal of the American Statistical Association*, 71(356), 864-869. <https://doi.org/10.1080/01621459.1976.10480960>
- Hamouda, A. (2013). An investigation of listening comprehension problems encountered by Saudi students in the EL listening classroom. *International Journal of Academic Research in Progressive Education and Development*, 2(2), 113-155.
- Hasançebi, B., Terzi, Y., & Küçük, Z. (2020). Madde güçlük indeksi ve madde ayırt edicilik indeksine dayalı çeldirici analizi [Distractor analysis based on item difficulty index and item discrimination index]. *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 10(1), 224-240. <https://doi.org/10.17714/gumusfenbil.615465>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132. <https://doi.org/10.2307/327317>
- Kahraman, A. (2015). Dinleme becerisi öğretimi ya da külkedisi sinderella [Teaching listening skill or the Orphan Cinderella]. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (45), 144-159. <https://dergipark.org.tr/en/pub/dpusbe/issue/4787/66070>
- Kemiksiz, Ö. (2017). Dinleme becerisi üzerine yazılan makalelerin değerlendirilmesi [Assessment of articles on listening skill]. *International Journal of Languages' Education and Teaching*, 5(1), 511-531. <https://dergipark.org.tr/tr/download/article-file/3674992>
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.
- Krashen, S. (1996). *Principles and practice in second language acquisition*. Pergamon Press.
- Levy, K. J. (1980). A Monte Carlo study of analysis of covariance under violations of the assumptions of normality and equal regression slopes. *Educational and Psychological Measurement*, 40(4), 835-840. <https://doi.org/10.1177/001316448004000404>
- Manurung, M. M. S., Purba, R., & Munthe, M. V. R. (2022). The effect of using LyricsTraining website on the students' listening skill of grade 8th. *Jurnal Pendidikan Dan Konseling*, 4(5), 5942-5950.
- Medina, S. L. (2002). Using music to enhance second language acquisition: From theory to practice. In J. Lalas & S. Lee (Eds.), *Language, literacy, and academic development for English language learners*. Pearson Education

- Meilinda, S. (2020). *An analysis of students perception on the use of Lyrics Training application to improve students' listening skills at sixth semester students of IAIN Salatiga in the Academic Year 2019/2020* [Graduation Paper, State Institute for Islamic Studies (IAIN) Salatiga]. <http://perpus.iainsalatiga.ac.id/lemari/fg/free/pdf/?file=http://perpus.iainsalatiga.ac.id/g/pdf/public/index.php/?pdf=8489/1/SISKA%20MEILINDA%20SKRIPSI%20pdf>
- Melanlıoğlu, D. (2013). Ortaokul öğrencileri için dinleme kaygısı ölçeğinin geçerlik ve güvenirlik çalışması [Reliability and validity of the listening anxiety scale for secondary school students]. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(11), 851-876. <https://www.ajindex.com/dosyalar/makale/acarindex-1423866546.pdf>
- Merzifonluoğlu, A. (2023). Promoting listening skills with an educational tool: VoScreen. *Journal of Educational Technology & Online Learning*, 6(4), 1027-1043. <https://doi.org/10.31681/jetol.1343391>
- MoNE. (2018). *İngilizce dersi öğretim programı ilkök ve ortaokul 2, 3, 4, 5, 6, 7 ve 8. Sınıflar* [The National English Language Curriculum for Primary and Lower Secondary Schools (Grades 2-8)]. <https://mufredat.meb.gov.tr/Dosyalar/201812411191321-%C4%B0NG%C4%B0L%C4%B0ZCE%20%C3%96%C4%99ERET%C4%B0M%20PROGRAMI%20Klas%C3%B6r%C3%BC.pdf>
- Mukhtorova, M., & Ilxomov, X. (2024). How to improve listening skills of both ESL and EFL students. *Qo'qon Universiteti Xabarnomasi*, 11, 84-86. <https://doi.org/10.54613/ku.v11i1.964>
- Nguyen, T. G. M. (2023). The impacts of lyrics training on improving listening skills for ESL students. *AsiaCALL Online Journal*, 14(1), 18-29. <https://doi.org/10.54855/acoj.231412>
- Nunan, D. (1999). *Second language teaching & learning*. Heinle & Heinle.
- Ocak, G., Küçükçınar, A., & Karakuyu, A. (2022). Ortaokul öğrencileri için İngilizce dil becerileri öz yeterlik ölçeği geliştirme çalışması [Developing English Language skills self efficacy scale for secondary school students]. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (40), 1-36. <https://doi.org/10.14520/adyusbd.1053460>
- OECD. (2021). Education at a Glance. OECD. <https://doi.org/10.1787/b35a14e5-en>
- Okcan, A. (2023). *Görev temelli dil öğretimi yaklaşımının 4. sınıf İngilizce dersi dinleme becerisine etkisinin incelenmesi* [Investigation of the effect of task-based language teaching approach on 4th grade English listening skills] (Thesis No. 837824) [Master's thesis, Nevşehir Hacı Bektaş Veli University]. Council of

- Higher Education Thesis Center.
<https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Orakcı, Ş. (2018). İngilizce Konuşma Kaygısı Ölçeği'nin geçerlik ve güvenilirlik çalışması [A validity and reliability study of "English Speaking Anxiety Scale"]. *İnönü Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, 5(9), 1-28. <https://doi.org/10.29129/inujse.375916>
- Paras, P. (2024). *The influence of using lyricstraining applications toward students' listening skills* (Publication No. 1911040166) [Doctoral dissertation, State Islamic University]. <https://repository.radenintan.ac.id/33916/1/PERPUS%20PUSAT%20BAB%201%20DAN%202.pdf>
- Polat, M. (2019). *Otantik videoların farklı İngilizce yeterlik düzeylerinde yabancı dilde dinleme becerisinin geliştirilmesi ve yabancı dilde dinleme kaygısı üzerindeki etkisi [The effects of authentic videos on foreign language listening skill development and foreign language listening anxiety at different levels of English proficiency]* (Thesis No. 553230) [Doctoral dissertation, Anadolu University]. Council of Higher Education Thesis Center. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Say Yağcıoğlu, S., & Çalışkan, M. (2024). The effect of games on listening and speaking anxiety in learning English. *Participatory Educational Research*, 11(4), 267-283. <https://doi.org/10.17275/per.24.60.11.4>
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences*. Lawrence Erlbaum Associates.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Allyn and Bacon.
- Turgut, M. F., & Baykul, Y. (2018). *Eğitimde ölçme ve değerlendirme [Measurement and Evaluation in Education]*. Pegem Akademi.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191-210. <https://doi.org/10.1017/S0261444807004338>
- Veloz Toapanda, M. E. (2023). *The use of LyricsTraining website to improve communicative skills* [Master's thesis, Universidad Técnica de Ambato]. <https://repositorio.uta.edu.ec/items/840a089c-5bcf-462a-8adf-6aa7c07a26b1>
- Yukselir, C., & Ozer, O. (2022). EFL learners' views on learner autonomy and foreign language anxiety. *Acuity: Journal of English Language Pedagogy, Literature and Culture*, 7(2), 248-263. <https://doi.org/10.35974/acuity.v7i2.2824>
- Zhang, X. (2013). Foreign language listening anxiety and listening performance. *System*, 41(1), 164-177. <https://doi.org/10.1016/j.system.2013.01.004>

Ethical and Author Declarations | Etik ve Yazar Beyanları

Graduate Thesis Information

This study is derived from the first author's 2025 master's thesis at Ordu University Graduate School of Social Sciences, supervised by the second author (thesis no: 926625). The thesis title is "The effect of the use of Web 2.0 tools in foreign language teaching on the listening achievement and anxiety of seventh grade students: A LyricsTraining example."

Authors' Contributions

This study is derived from the master's thesis completed by the first author under the supervision of the second author. The first author conducted the research, performed data curation, and drafted the original article. The second author supervised the study and provided critical feedback on the findings. Both authors contributed to the final version of the article and approved it for publication.

Conflict of Interest

The authors declare that they have no competing interests relevant to the content of this article.

Ethical Approve

The planning and conduct of this research was approved by the Ordu University Rectorate Social and Humanities Research Ethics Committee decision dated 01/03/2023 and numbered 2023-107.

Lisansüstü Tez Bilgileri

Bu çalışma, birinci yazarın 2025 yılında Ordu Üniversitesi Sosyal Bilimler Enstitüsü'nde ikinci yazar danışmanlığında tamamladığı yüksek lisans tezinden türetilmiştir (tez no: 926625). Tez başlığı: "Yabancı dil öğretiminde Web 2.0 araçları kullanımının yedinci sınıf öğrencilerinin dinleme başarılarına ve kaygılarına etkisi: LyricsTraining örneği "

Yazarların Katkı Düzeyleri

Bu çalışma birinci yazar tarafından ikinci yazarın danışmanlığında tamamlanan yüksek lisans tezinden üretilmiştir. Birinci yazar araştırmayı yürüterek veri düzenlemesini gerçekleştirmiş ve makalenin özgün taslağını yazmıştır. İkinci yazar çalışmayı denetlemiş ve bulgulara ilişkin eleştirel geri bildirim sağlamıştır. Her iki yazar da makalenin son haline katkıda bulunmuş ve yayımlanmasını onaylamıştır.

Çıkar Çatışması Beyanı

Yazarların, bu makalenin içeriğiyle ilgili beyan edilecek herhangi bir çıkar çatışması bulunmamaktadır.

Etik Onay

Bu araştırmanın planlanması ve yürütülmesi, Ordu Üniversitesi Rektörlüğü Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu'nun 01/03/2023 tarih ve 2023-107 nolu kararı ile uygun bulunmuştur

This study has been evaluated under double-blind peer review and verified to be free of plagiarism using iThenticate software.

Bu çalışma, çift taraflı kör hakemlik kapsamında değerlendirilmiş ve iThenticate yazılımı kullanılarak intihal içermediği teyit edilmiştir.

The copyrights of the studies published in our journal belong to our journal and are published as open access under a CC-BY-NC-ND license.

Dergimizde yayımlanan çalışmaların telif hakları dergimizde olup, CC-BY-NC-ND lisansı ile açık erişim olarak yayımlanmaktadır.

Ethical disclosure | Etik bildirim: ebfd@ankara.edu.tr