



(ISSN: 2587-0238)

Çenberci, S. & Kalkanoğlu, B. (2023). Turkish Adaptation of Singing Voice Development Measure (SVDM) and Examination of Children's Use of Singing Voice on Different Variables, *International Journal of Education Technology and Scientific Researches*, 8(25), 2949-2982.

DOI: <https://dx.doi.org/10.35826/ijetsar.715>

Article Type (Makale Türü): Research Article

---

## **TURKISH ADAPTATION OF SINGING VOICE DEVELOPMENT MEASURE (SVDM) AND EXAMINATION OF CHILDREN'S USE OF SINGING VOICE ON DIFFERENT VARIABLES**

**Senim ÇENBERCİ**

Research Assistant Dr., Gazi University, Ankara, Türkiye, [scenberci@hotmail.com](mailto:scenberci@hotmail.com)  
ORCID: 0000-0002-5692-2595

**Burcu KALKANOĞLU**

Assoc. Prof. Dr., Trabzon University, Trabzon, Türkiye, [burcukalkanoglu@trabzon.edu.tr](mailto:burcukalkanoglu@trabzon.edu.tr)  
ORCID: 0000-0003-3696-5907

Received: 18.06.2023

Accepted: 10.11.2023

Published: 05.12.2023

### **ABSTRACT**

This study aimed to adapt the "Singing Voice Development Measure (SVDM)" developed by Rutkowski into Turkish and examine the children's singing voices in terms of different variables, such as age, gender, and the use of songs with/without words. For this purpose, after obtaining the necessary permissions, two independent translators translated the measurement tool. The translation was finalized after being submitted to expert opinions within the scope of the linguistic equivalence study. The research group consisted of 253 children who attended between kindergarten and 3rd grade in a primary school in the Çankaya district of Ankara in the 2022/2023 academic year. The analyses carried out according to the Rasch model revealed that the measurement tool works in line with the purpose, contains positive findings regarding its validity, and can be used for the purpose. Regarding the reliability of the measurement tool, it was determined that the inter-rater and intra-rater reliability values were high. These findings show that SVDM is a valid and reliable measurement tool in Turkish culture. When the singing voice usage of Turkish children is examined in terms of different variables, it is seen that the use of singing voice increases as the grade level increases. It has been determined that girls are relatively more advanced in the use of singing voice than boys, and compared to songs with words, songs without words provide an advantage in the use of the singing voice of children. Incorporating the SVDM into our education system for use in general music education classes is recommended. It is also recommended to conduct large-sample screening studies at different grade levels to measure the development of children's singing voices in terms of different variables to eliminate and improve the current deficiencies in our music education system.

**Keywords:** Singing voice development measure (SVDM), singing, general music education, validity, reliability.

## INTRODUCTION

General music education refers to a type of music education designed for all students without discrimination, intending to develop basic musical skills, knowledge, and understanding. General music education, a part of primary and secondary school curricula, is carried out today by music educators and classroom teachers worldwide. It includes learning objectives for singing, playing, creativity, and interacting with music (Abril, 2016, p. 5). Among these learning objectives, singing is the most fundamental activity in general music education classes. The most important duty of the music educator is to teach children to sing and encourage children's desire to sing from the first years of general music education (Mizener, 2008, p. 23). To this end, the music educator needs to understand the developmental nature of the child's singing voice and develop a teaching program according to individual needs.

The foundations of singing skills are laid in the womb. The mother's speaking and singing voice is transmitted to the fetus through sound waves. The auditory and emotional experiences of the developing fetus in the last months of pregnancy constitute the primary source of musical interaction and singing skills (Welch, 2016, p. 442). Musical development begins before birth and is shaped in line with the quality of the musical environment offered to babies/children after birth. Babies are born with the potential to use both speaking and singing voices. Adults constantly model the speaking voice, so babies can quickly improve their speaking skills. Unfortunately, singing skills lag behind the development of speaking skills. The main reason is that babies are not sufficiently exposed to the singing voice in the environment where the native language is spoken. Thus, babies are more likely to use speaking voice rather than discover singing voice (Gordon, 2007, pp. 18-19). In this context, it is important to provide babies with a musical environment at birth where they can learn the musical language in the informal ways they learn their mother tongue. Just as children with a specific vocabulary in their native language acquire speaking skills, children with a specific musical vocabulary can also acquire a singing voice from a very young age. According to Gordon (2013, p. 142), just as the speaking voice is learned by listening to and imitating models in the environment, the singing voice is learned similarly by listening to and imitating a good model based on live interactions.

Ideally, a rich musical vocabulary should be constituted before starting general music education, and audition skills should be developed between the ages 0-6. However, every child cannot have access to informal music learning from an early age. In this case, it is seen that children who have started kindergarten and primary school cannot discover the singing voice or have weak intonation (Gordon, 2007, p. 51). Welch (1986, p. 300) lists the problems that children may encounter regarding singing in general music education classes as follows:

1. Children may have difficulty hearing their voices and responses in group singing activities,
2. Difficulties may arise due to the lack of opportunity to sing individually and receive individual help,
3. Before starting the school music repertoire, children need to work in a limited vocal register and gain confidence in that register. Difficulties may arise due to the absence of this practice.

4. Children may need to study music and lyrics separately from each other (Especially for some children, there is a need to study only music with natural syllables, without the complexity of the language).
5. Children need to learn how to use their voice more effectively by practicing issues like breathing control.

Individuals who have not been able to use their singing voice have been described with various titles in the literature. These include titles such as monotone, uncertain-singer, out-of-tune singer, poor pitch singer, problem singer, and inaccurate singer (Aaron, 1990, p. 4). Most of these titles do not emphasize the developmental aspects of singing skills. The inability to use the singing voice is not an innate disability associated with the lack of the ability to distinguish sounds but is the product of various environmental factors (Welch, 1979, p. 56). The general opinion is that singing is a teachable and learnable skill (Welch, 1985, 1986; Gould, 1968, 1969; Smith, 2006; Mizener, 2008; Rutkowski & Miller, 2003c). Welch (1985, p. 3) expresses the developmental nature of singing skills with the following words: "For children, singing in tune is not simply a question of 'can/cannot' but rather a hierarchy or continuum of skill with several degrees of competence." In this context, children who cannot use their singing voice (out-of-tune singers) are considered developing singers who can benefit from training (Leighton & Lamont, 2006, p. 311).

Gould (1968, p. 8) emphasizes that the biggest obstacle to young children's singing stems from the psychological obstacles caused by the social environment:

"Actually, it is doubtful if there is such a thing as a non-singing child at the ages of nursery school and kindergarten. They all sing. Perhaps some of them do not sing the way the teacher wants them to; but they all do sing. It is only after problem singers have experienced unfavorable reactions to their singing efforts from peers, teachers, or parents that psychological inhibitions develop".

Children who cannot discover or learn to use the singing voice from an early age move away from music and become alienated from it as they grow up (Gould, 1969, p. 11). It is critical to determine the stage of children's singing skills and individual needs in terms of applying remedial techniques and compensatory guidance accordingly. At this point, it is necessary to use the correct measurement tool for the purpose. Salvador (2010, p. 43) emphasizes that when designing or adapting a measurement tool to measure singing behavior, attention should be paid to issues, such as whether the measurement tool is a valid tool to measure singing skills, the measurement tool provides acceptable reliability values, the criteria are clear and understandable, the measurement tool can be included in the classroom environment, and it can provide a diagnosis to determine the needs of each child.

When the international literature is examined, it is seen that the measurement tools for singing behavior are divided into two: tools for measuring the development of a singing voice (Rutkowski, 1990, 1997) and tools for measuring singing accuracy (Boardman, 1964; Dittmore, 1968; Hale, 1976; Newlin, 2003; Hornbach&Taggart, 2005; Wise&Sloboda, 2008). Nicols (2017, p. 14) also emphasizes that tools for singing should be designed and

used for the feature to be measured. If the aim is to measure the development of the singing voice of the students, tools that measure the developmental characteristics of the vocal register and singing behavior, such as the use of singing voice or speaking voice, should be used; if the aim is to measure the vocalization of certain pitches with a good intonation or singing in tune, tools that measure the accuracy of singing behavior should be used. The development of singing voice and vocal accuracy are separate but interrelated skills. Rutkowski (2015) revealed a positive correlation between kindergarten and first-grade children's use of singing voice and their vocal accuracy. The results obtained from this study suggest that focusing on how well children who cannot yet use the vocal register can sing (vocal accuracy, singing in tune) may not be beneficial for developing singing skills. In this regard, it is crucial to determine the singing voice development (vocal register) before measurements of children's vocal accuracy are made.

Two critical studies present developmental models by dividing the acquisition of a singing voice into stages (Welch, 2002; Rutkowski, 1997). Welch (2002) developed the "vocal pitch-matching development" model of four stages. According to this model, in the first stage, the lyrics are in the foreground rather than the melody; the melody is mostly sung as if chanting, and a limited vocal register is used. In the second stage, awareness that changes in pitch can be controlled increases, the melodic line can be followed in general terms, and the vocal register used expands. In the third stage, melodic lines and intervals mostly gain accuracy; however, some changes in tonality may occur. In the fourth stage, there are no significant errors in the melodic line or pitches, and the singing voice becomes more accurate.

On the other hand, Rutkowski systematically evaluated children's singing behavior for more than fifteen years and presented a nine-stage model (Welch, 2016, p. 446). According to this model, children progress through nine stages, from using the speaking voice range to using a limited vocal range, the initial vocal range, and the expanded vocal range. The Singing Voice Development Measure (SVDM) is also based on this model. SVDM is a practical, understandable, reliable measurement tool for general music education classes. It can be used to evaluate the performance of songs whose vocal range consists of intervals D 3-B flat 3 and above. In addition, it evaluates the singing voice from a developmental perspective and provides diagnostic information about what children need to increase their achievements (Salvador, 2010, pp. 44-45).

When the national literature is reviewed, it is seen that measurement tools designed to measure singing skills primarily focus on accuracy in singing and the development of singing behaviors from various aspects rather than the developmental characteristics of singing behavior (Bilen, 1995; Uçal, 2003; Gürgen, 2007; Uysal, 2004; Karaönçel, 2021). Among these tools, only Değer (2012) developed the "Behavior Development Scale in the Use of Singing Voice" to measure the development of students' singing voice and demonstrated its validity and reliability. This scale is designed to determine the development of the singing voice, from using the speaking voice to various vocal registers consisting of sixth, eighth, and above intervals. When the literature was examined, its validity and reliability were discussed in detail in a large sample, and it was determined that there was a need for a measurement tool that evaluated singing voice from a developmental perspective. Rutkowski's "Singing

Voice Development Measure" is a measurement tool whose validity and reliability have been demonstrated and has a wide usage rate in the international literature (Tavani, 1998; Jaffurs, 2000; Kim, 2000; Guerrini, 2002, 2004, 2006; Doneski, 2005; Mang, 2006; Vande Wege, 2005, Steenwyk, 2004; Hall, 2005; Tu, 2009; Kendal, 2018; Pereira & Rodrigues, 2019). Based on this, the current study aimed to adapt the "Singing Voice Development Measure" into Turkish, test its validity and reliability as a performance-based measurement tool, and evaluate the singing voice development in terms of various variables. For this purpose, the following questions were addressed respectively:

1. What are the validity and reliability values for the with-words and without-words (natural syllable) evaluation criteria of SVDM?
2. Do students' performances on with-words and without-words (natural syllable) evaluation criteria show a significant difference according to their grade level?
3. Do students' performances on with-words and without-words (natural syllable) evaluation criteria differ significantly according to their gender?
4. Is there a statistically significant relationship between students' performance on with-words and without-words (natural syllable) evaluation criteria?

## METHOD

### *Research Group*

The research group consisted of 253 children who attended classes in kindergarten and 3rd grade in a primary school in the Çankaya district of Ankara in the 2022/2023 academic year. Participation in the research was based on volunteering. Consent forms prescribed by the Gazi University Ethics Commission (decision dated 10.03.2023 and numbered 2023 – 248) were collected from parents. The distribution of participants according to gender and grade level is presented in Table 1 below.

**Table 1:** Distribution of The Participants by Gender and Grade Level.

Grade Level	Boy	Girl	Total
Kindergarten	7	11	18
Grade-1	55	35	90
Grade-2	42	48	90
Grade-3	26	29	55
Total	130	123	253

### *Data Collection Tool*

#### *Singing Voice Development Measure (SVDM)*

Rutkowski introduced this scale into the literature (1990, 1997) as a result of many years of development and revision studies to measure children's use of singing voice. The questions comprise eight patterns in the D 3-B flat 3 range. Before the application, these patterns are studied with the whole class for a while as preliminary

preparation. The questions are administered to each child individually in a separate classroom environment following preliminary preparation. An instrument such as a piano/organ/tone bell, which is necessary to play the patterns, and a recording device so that raters can evaluate should be placed in this class. The application procedure takes place in the following steps;

1. The first pattern is played on the piano (e.g., A-F-D).
2. Then the teacher vocalizes the same pattern using natural syllables (bum and pam.) or words. Accompaniment is not used at this stage.
3. The child is asked to imitate the pattern. There is no accompaniment at this stage, either.
4. The first three stages are repeated for each pattern.
5. There is no waiting time between steps.
6. One-half of the group vocalizes the patterns first with natural syllables and then with words; the other half vocalizes first with words and then with natural syllables (Rutkowski & Miller, 2003, pp. 6-7).



**Figure 1:** Patterns of Singing Voice Development Measures.

**Table 2:** Rating Scale (Rutkowski & Miller, 2003, p. 8).

1	<b>“Pre-singer”</b> does not sing but chants the song text.
1,5	<b>“Inconsistent Speaking Range Singer”</b> sometimes chants, sometimes sustains tones, and exhibits some sensitivity to pitch but remains in the speaking voice range (usually A2 to C3).
2	<b>“Speaking Range Singer”</b> sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range (usually A2 to C3).
2,5	<b>“Inconsistent Limited Range Singer”</b> wavers between speaking and singing voice and uses a limited range when in singing voice (usually up to F3).
3	<b>“Limited Range Singer”</b> exhibits consistent use of limited singing range (usually D3 to F3).
3,5	<b>“Inconsistent Initial Range Singer”</b> sometimes only exhibits the use of limited singing range, but other times exhibits the use of initial singing range (usually D3 to A3).
4	<b>“Initial Range Singer”</b> exhibits consistent use of the initial singing range (usually D3 to A3).
4,5	<b>“Inconsistent Singer”</b> sometimes only exhibits the use of the initial singing range, but other times exhibits the use of extended singing range (sings beyond the register lift: B3-flat and above).
5	<b>“Singer”</b> the uses of consistent extended singing range (sings beyond the register lift: B3-flat and above).

### **Data Analysis**

The data analysis of this research consists of two stages. In the first stage, the multi-facet Rasch measurement model was calculated to test the measurement tool's validity. The research has five facets: evaluation (with-words and without-words), raters, student, class, and gender.

In the second stage, One-way analysis of variance and unrelated t-tests were used to examine whether students' average scores on with-words and without-words (natural syllable) performances differed according to class and gender variables. The Pearson product-moment correlation coefficient determined whether there was a significant relationship between the students' performances on with-words and without-words (natural syllable) criteria.

### **Adaptation Process: Translation of Scale Items into Turkish**

After the adaptation permission was obtained from Joanne Rutkowski, the scale's English-Turkish translation process started. The original form was translated into Turkish by an independent sworn translator and researcher fluent in both languages. The expressions obtained from the first translation were compared, and common expressions were determined. The expert opinion form, which includes the items in the original language and the common expressions obtained as a result of two different translations, was submitted to the opinions of 6 faculty members working in Gazi University, Gazi Education Faculty, Music Education Department and English Education Department within the scope of linguistic equivalence study. In line with the suggestions obtained from the expert opinions, necessary arrangements were made for the Turkish items, and English-to-Turkish translation took its form.

### **FINDINGS**

#### ***Findings and Interpretation Regarding the Question " What Are the Validity and Reliability Values for the With-Words and Without-Words (Natural Syllable) Evaluation Criteria of SVDM?"***

In this research question, measurement values for both criteria were examined as evidence of validity. Students' performance measurements on with-words and without-words criteria were analyzed according to the multi-facet Rasch measurement model. The values of the criteria are given in Table 3.

**Table 3.** Measurement Report of With-Words and Without-Words Criteria Criteria Surface

Criterion	Observed Average	Adjusted Average	Logit Value	Standard Error	In Fit	Out of Fit
With-words	3,196	3,243	-0,23	0,07	0,98	0,98
Without-words	3,309	3,424	0,23	0,07	1,00	0,97
Mean	3,252	3,333	0,00	0,07	0,99	0,98
S (Universe)	0,056	0,090	0,23	0,00	0,01	0,01
S (Sample)	0,080	0,128	0,33	0,00	0,01	0,01
Model, Universe	RMSE= 0,07 Corrected S.E.= 0,22 Separation Rate = 3,13 Separation Index = 4,51 Reliability= 0,91					
Model, Sample	RMSE= 0,07 Corrected S.E.= 0,32 Separation Rate= 4,54 Separation Index = 6,39 Reliability = 0,95					
Model, Chi-square (fixed) :	21,6 df= 1 p= 0,00					
S: Standard deviation, RMSE: Square root of mean square error, S.E.: Standard error, df: Degrees of Freedom						

When Table 3 is examined, it is seen that the logit values of the with-words and without-words criteria are -0.23 and 0.23, and the criteria have a logit change of 0.46 [0.23-(-0.23)]. The mean of logit values for the criteria is 0.00, and the standard deviation is 0.32. According to the group, these values indicate that the without-words evaluation was relatively easy, and the with-words evaluation was relatively difficult. When the criterion surface values are examined, the in-fit values are 0.98 and 1.00, while the out-of-fit values are 0.98 and 0.97. The mean of in-fit values is 0.99, and the mean of out-of-fit values is 0.98. The fact that the in-fit and out-of-fit values are close to 1 indicates that the model data fit is good. The separation index of the criterion surface is 6.39. The separation index above 2.00 shows that the criteria can be applied to individuals at different levels and are functional (Nakamura, 2002). The high separation rate and reliability index values indicate that the difficulty levels of the criteria are different. When the chi-square value was examined, it was determined that this difference was statistically significant ( $\chi^2(sd) = 21,6 (1), p = 0,00 < 0,01$ ). This shows that there is a significant difference between the criterion difficulties. The logit value of with-words and without-words measures, in-fit and out-of-fit values, and discrimination index show that the tool works in line with its purpose, contains positive findings regarding its validity, and can be used for the purpose.

Inter-rater reliability was examined as evidence of reliability for with-words and without-words criteria. Inter-rater reliability was calculated using the Krippendorff alpha coefficient (Hayes & Krippendorff, 2007). According to the findings, the reliability coefficient was obtained as  $\alpha = .99$  for the with-words criterion and  $\alpha = .98$  for the without-words criterion. These findings show that inter-rater reliability is high. The Pearson Correlation Coefficient determined intra-rater reliability values. Accordingly, the correlation coefficient between the first rater's evaluations at two different times was  $r = .94$  for the with-words criterion and  $r = .98$  for the without-words criterion. The correlation coefficient between the second rater's evaluations was determined as  $r = .96$  for the with-words criterion and  $r = .99$  for the without-words criterion. These values show that also the intra-rater reliability is quite high.



**Findings and Interpretation Regarding the Question, "Do Students' Performances on With-Words and Without-Words (Natural Syllable) Evaluation Criteria Show a Significant Difference According to Their Grade Level?"**

Whether students' performance scores on with-words and without-words criteria show a significant difference according to their grade level; since the scores of the test variable were normally distributed across grade levels, the scores were equally spaced, and the variance values of the groups were homogeneous, it was analyzed with one-way analysis of variance. Analyzes were made on 253 children. The findings obtained from the analysis are given in Table 4 below.

**Table 4.** Comparison of Students' Performances on With-Words and Without-Words Criteria According to Grade Levels Using Single Factor Analysis of Variance

Dimension	Category	N	Mean	SD	F	P	Difference (Tukey)	Eta square
With-words	Kindergarten	18	2,35	1,05	6,047	,001	Kindergarten-G.2 Kindergarten-G.3	0,28
	Grade 1	90	3,03	1,19				
	Grade 2	90	3,45	1,27				
	Grade 3	55	3,53	1,20				
	<b>Total</b>	253	3,24	1,25				
Without-words	Kindergarten	18	2,44	1,20	4,981	,002	Kindergarten-G.2 Kindergarten-G.3	0.26
	Grade 1	90	3,18	1,22				
	Grade 2	90	3,58	1,30				
	Grade 3	55	3,52	1,29				
	<b>Total</b>	253	3,34	1,29				

The performances of students attending at different grades on the with-words criterion were compared according to grade levels, and the differences between the groups were found to be statistically significant ( $F(3, 249) = 6.047, p < .05$ ). In other words, it has been determined that with-word evaluation varies according to grade levels. When the results of the multiple comparison Tukey (HSD) test, which shows which pairwise groups the differences originate from, were examined, a relative difference was found between the with-words evaluation scores of kindergarten ( $M = 2.35; SD = 1.05$ ) and first grade ( $M = 3.03; SD = 1.19$ ) students; however, this difference is not statistically significant. It was determined that there was a significant difference between the performances of kindergarten and second grade ( $M = 3.45; SD = 1.27$ ) and kindergarten and third grade ( $M = 3.53; SD = 1.20$ ) students regarding the with-words criterion ( $p < .05$ ).

When the students' performances on the without-words criterion were compared according to their grade levels, a statistically significant difference was found between the groups ( $F(3, 249) = 4.981, p < .05$ ). In other words, performances on without-words criteria varied according to grade level. Multiple comparison Tukey (HSD) test was performed to determine which pairwise group the differences originated from. When the results were examined, no significant difference was found between the without-word evaluations of kindergarten ( $M = 2.44; SD = 1.21$ ) and first-grade ( $M = 3.18; SD = 1.22$ ) students. It was determined that there was a relative difference between the without-word evaluation scores of primary school second ( $M = 3.58; SD = 1.30$ ) and third-grade ( $M =$

3.52; SD=1.29) students ( $p < .05$ ). Although it is seen that third-grade performances in the without-word criterion decrease relatively compared to second grade, this difference is not statistically significant. In addition, when the with-words and without-words performance averages are examined according to grade levels, it is seen that the students' performances on the without-words criterion are relatively higher than their performances on the with-words criterion.

Another criterion that shows whether the difference between the groups' data is significant is the effect size (Kılıç, 2014). Since the effect size value calculated for both evaluations is greater than 0.8, it is defined as a strong effect (Cohen, 1988). It can be seen from here that the effect sizes are at similar levels.

**Findings and Interpretation Regarding the Question "Do Students' Performances on With-Words and Without-Words (Natural Syllable) Evaluation Criteria Differ Significantly According to Their Gender?"**

Whether the students' performance scores on with-words and without-words criteria showed a significant difference according to their gender was examined with the unrelated t-test since the scores of the groups were normally distributed, and the scores were equally spaced. Analyzes were made on 253 children. The findings obtained from the analysis are given in Table 5.

**Table 5.** Comparison of Students' Performances on With-Words and Without-Words Criteria According to Their Gender with Unrelated T-Test

	Gender	N	Mean	SD	DF	t	p
With-words	girl	123	3,53	1,19	251	3,600	,000
	boy	130	2,97	1,25			
Without-words	girl	123	3,62	1,25	251	2,721	,001
	boy	130	3,08	1,28			

A significant difference was found in the students' performances on with-words and without-words criteria according to their gender ( $t_{(253)} = 3,600; p < 0.5$ ). The performance average of female students regarding with-words evaluation (Mean = 3.53; SD = 1.19) was higher than the performance average of male students regarding with-words evaluation (Mean = 2.97; SD = 1.25). Likewise, a significant difference was found between the performance scores on the without-word criterion according to gender ( $t_{(253)} = 2,721; p < 0.5$ ). Female students' mean for without-words evaluation (Mean = 3.62; SD = 1.25) was again higher than male students' mean for without-words evaluation (Mean = 3.08; SD = 1.28). These results show that girls' with-words and without-words performances are better than boys. In other words, it can be said that the student's gender is effective in the performance of the criteria.

**Findings and Interpretation Regarding the Question "Is there a statistically significant relationship between students' performance on with-words and without-words (natural syllable) evaluation criteria?"**

Students' performances on with-words and without-words criteria were calculated by taking the average of the scores given by two experts. The relationship between performances on with-words and without-words criteria according to the average scores of two experts, since the scores of both variables were normally distributed and

equally spaced, was determined by the Pearson Product Moment Correlation Coefficient. It was observed that there was a high level of positive relationship between the variables ( $r = .86$   $p < 0.05$ ). Based on this finding, it can be said that as students' with-words performances increase, their without-words performances also increase significantly, or as students' with-words performances decrease, their without-words performances decrease significantly.

## **CONCLUSION and DISCUSSION**

According to the findings obtained from the analyses made according to the Rasch model, the logit value of with-words and without-words measures was -0.23 and 0.23, and the logit change value was 0.46 [0.23-(-0.23)]; The in-fit mean is  $M = 0.99$  and the out-of-fit mean is  $M = 0.98$ ; The separation index was obtained as 6.39. These findings suggest that the measurement tool works for its purpose, contains positive results regarding its validity, and can be used for its purpose. Studies on the content validity of the measurement tool in its original culture have proven that it is a valid instrument for measuring the intended feature (Rutkowski, 1990, p. 92). The results of the current study reveal the validity of the measurement tool in Turkish culture.

Regarding the reliability of the measurement tool, inter-rater and intra-rater reliability values were calculated. Accordingly, inter-rater reliability values were obtained as  $\alpha = .99$  for the with-words criterion and  $\alpha = .98$  for the without-words criterion. Intra-rater reliability values for with-words and without-words criteria were obtained as  $r = .94$  and  $.98$  for the first rater and  $r = .96$  and  $.99$  for the second rater. In the original reliability study of the third version of the Singing Voice Development Measure, the Pearson Product Moment Correlation Coefficient was used to determine inter-rater and intra-rater reliability values. Accordingly, it was determined that the inter-rater reliability coefficient of the measurement tool varied between  $.80$  and  $.85$ , and the intra-rater reliability coefficient varied between  $.82$  and  $.93$ . These values indicate the reliability of the measurement tool within its original culture (Rutkowski, 1997, pp. 206-207). The results of the current adaptation study show that the measurement tool also provides high-reliability values in Turkish culture.

Pereira and Rodrigues (2019) discussed the validity and reliability of the Singing Voice Development Measure in Portuguese culture. The results of expert opinions regarding content validity showed that the measurement tool was valid for Portuguese children. Although the frequency of the criteria of Pre-Singer (1), Inconsistent, speaking range singer (1.5), and speaking range singer (2) is relatively low, the overall results indicate the validity of the measurement tool in the Portuguese culture. Inter-rater reliability coefficients were obtained as  $.95$  for both with-words and without-words criteria. This study revealed the validity and reliability of the measurement tool in Portuguese culture. Chuang (2012) adapted the measurement tool to Chinese culture and expanded the nine criteria in the original measurement tool to obtain a structure with eleven criteria. The results of the current study revealed that the measurement tool also produced valid and reliable results on Turkish children with its nine-criterion structure.

When the sample's singing voice usage was examined according to the grade level variable, it was determined that the scores of the 3rd graders in the without-words criterion were relatively lower than the scores of the 2nd graders; however, this difference is not statistically significant. In general, as the grade level increases, the level of singing voice usage tends to increase. This finding aligns with the studies in the literature suggesting that biological age impacts the use of singing voice and singing accuracy (Davies & Roberts, 1975; Goetze, 1985; Rutkowski & Miller, 2003a). However, there are also studies suggesting that biological age does not significantly affect the development of singing voice and singing skills (Mang, 2006; Levinowitz et al., 1998). These contradictory findings in the literature may be due to sample size, sample selection, and cultural differences. The results of Rutkowski & Miller's (2003a) long-term study on primary school children revealed that children made more significant gains in singing skills after the first grade. The same study's results suggest that with small group and individual singing strategies implemented within the scope of general music education, most children can use at least the initial vocal range by the end of the fifth grade.

When the gender variable was considered, it was determined that girls received higher scores on the Singing Voice Development Measure than boys. This difference suggests that girls have an advantage in using singing voice compared to boys. When the literature is examined, it is seen that there are studies indicating that gender affects the use of singing voice and singing accuracy in favor of girls (Davies & Roberts, 1975; Goetze, 1985; Rutkowski, 1986; Guerrini, 2002, 2004; Mang, 2006). The current study is parallel to the findings of these studies. In her doctoral thesis, Rutkowski (1986, p. 243) found that although there was no statistically significant difference in the use of singing voice and developmental music aptitude in preschool children, girls achieved higher scores. According to this study, the difference between the two genders was slight. Contrary to these studies, some do not find evidence that gender affects singing skills (Smale, 1987; Welch et al., 1995; Welch et al., 1997; Cooper, 1995). Welch et al.'s (1997) three-year long-term study on the 5-7 age group suggests that the vocal pitch-matching skills of girls and boys are homogeneous, unlike previous studies. A difference in favor of girls regarding singing was obtained only in the 7-year-old group.

A highly significant relationship was determined between the with-words and without-words criteria of the measurement tool. This finding shows that as students' with-words performance increases, their without-words performance also increases significantly, or as students' with-words performance decreases, their without-words performance decreases significantly. When the performance averages regarding with-words and without-words criteria were examined, it was determined that the students' performances regarding the without-words criterion were relatively higher than their performances regarding the with-words criterion. This finding suggests that children between kindergarten and third grade of primary school sing songs more easily and clearly, regardless of their lyrics. Although there is no definitive consensus in the literature on the effectiveness of teaching songs with words or without words, the results of significant studies suggest that children can use the singing voice better and sing more accurately when song teaching occurs independently of the lyrics (Goetze, 1985; Welch et al., 1995; Rutkowski, 2013; Rutkowski & Miller, 2003b). The results of this study are in line with the relevant literature. Rutkowski (2013) found that preschool children significantly improved their use of singing

voice when they learned songs with natural syllables in an informal guidance-based music learning environment. Welch et al. (1995), as a result of their study on a five-year-old group, obtained findings revealing that children gave more precise vocal responses to musical elements that were not masked by words. They suggested that for young developing singers, lyrics may come before musical elements; therefore, young children should be taught music and lyrics separately. The results Kendal (2018) obtained differ between the development of a singing voice and ensuring pitch accuracy. Accordingly, she suggested that presenting songs primarily without words could support preschool children's singing voice development while presenting songs with words could support preschool children's pitch accuracy. Contrary to this evidence, Levinowitz (1987) and Smale (1987) could not find that using songs with words or without words significantly affected children's singing achievements.

The results of the current study revealed that the "Singing Voice Development Measure," which was put forward by Rutkowski as a result of many years of development and revision studies, produced valid and reliable results on Turkish children between kindergarten and 3rd grade of primary school. The use of this measurement tool, whose validity and reliability have been proven in Turkish by classroom teachers or music teachers teaching general music education courses in kindergarten and primary school, will significantly contribute to determining children's individual needs and designing teaching programs accordingly. In addition, the measurement tool, whose validity and reliability in Turkish has been handled, will fill an essential gap in the national literature and serve as a source for future research.

## **SUGGESTIONS**

It is recommended that studies be carried out on the use and dissemination of the Singing Voice Development Measure, which has been adapted into Turkish and whose validity and reliability have been demonstrated, especially in pre-school and primary school classes, and to be included as a measurement tool in existing music education practices. This study is limited to the sample comprising 253 children who attended kindergarten and 3rd-grade primary school. The validity of the measurement tool can be tested at a broader grade level by including all levels of preschool and primary school. In addition, it is recommended that large-sample screening studies at different educational levels be conducted to measure the development of Turkish children's singing voice in terms of different variables. Such studies to determine the situation will guide the elimination and improvement of the current deficiencies in our music education system.

## **ETHICAL TEXT**

In this article, the journal writing rules, publication principles, research and publication ethics, and journal ethical rules were followed. The responsibility belongs to the author (s) for any violations that may arise regarding the article. The ethics committee permission for the article was obtained by the Gazi University Ethics Commission with the decision dated 10.03.2023 and numbered 2023 - 248.

**Author(s) Contribution Rate:** Both authors contributed equally to the article. The authors would like to thank the Gazi University Academic Writing Application and Research Center for proofreading the article.

**REFERENCES**

- Aaron, J. C. (1990). *The effects of vocal coordination instruction on the pitch accuracy, range, pitch discrimination, and tonal memory of inaccurate singers* [Unpublished doctoral dissertation]. University of Iowa.
- Abril, C. R. (2016). Untangling general music education: concept, aims, and practice. In C. R. Abril & B. M. Gault (Eds.). *Teaching general music: Approaches, issues, and viewpoints*. (pp. 5-22). Oxford University Press.
- Bilen, S. (1995). *İşbirlikli Öğrenmenin Müzik Öğretimi ve Güdusel Süreçler Üzerindeki Etkileri* [Yayımlanmamış Doktora tezi]. Dokuz Eylül Üniversitesi.
- Boardman, E. L. (1964). *An investigation of the effect of preschool training on the development of vocal accuracy in young children* [Unpublished doctoral dissertation]. University of Illinois.
- Chuang, M. J. (2012). Construction of a children's singing voice measure. *Problems in Music Pedagogy*, 11, 93-101.
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale:Lawrence Erlbaum Associates.
- Cooper, N. A. (1995). Children's singing accuracy as a function of grade level, gender, and individual versus unison singing. *Journal of Research in Music Education*, 43(3), 222-231. <https://doi.org/10.2307/3345637>
- Davies, A. D. M., & Roberts, E. (1975). Poor Pitch singing: a survey of its incidence in school children. *Psychology of Music*, 3(2), 24-36. <https://doi.org/10.1177/030573567532004>
- Değer, A. Ç. (2012). *Çocuk korolarının eğitiminde bir yaklaşım olarak eğitsel oyun kullanımının öğrencilerin müziksel erişimi düzeylerine etkisi* [Yayımlanmamış Doktora tezi]. Gazi Üniversitesi.
- Dittemore, E. E. (1968). *An Investigation Of Some Musical Capabilities Of Elementary School Students* [Unpublished doctoral dissertation]. University of Iowa.
- Doneski, S. M. (2005). *The effects of wait time on the tonal pattern performance accuracy of second and fourth - grade students*. [Unpublished doctoral dissertation]. University of Hartford.
- Goetze, M. (1985). *Factors affecting accuracy in children's singing (music, Voice)*. [Unpublished doctoral dissertation]. University of Colorado.
- Gordon, E. E. (2007). *Awakening new-borns, children, and adults to the world of audiation*. GIA.
- Gordon, E. E. (2013). *Music Learning Theory for new-born and young children*. GIA.
- Gould, A. O. (1968). *Finding and learning to use the singing voice: a manual for teachers*. Eric Pub.
- Gould, A. O. (1969). Developing Specialized Programs for Singing in the Elementary School. *Bulletin of the Council for Research in Music Education*, 17, 9-22.
- Guerrini, S. C. (2002). *The acquisition and assessment of the developing singing voice among elementary students* [Unpublished doctoral dissertation]. Temple University
- Guerrini, S. C. (2004). The relationship of vocal accuracy, gender, and music aptitude among elementary students. *Visions of Research in Music Education*, 4(1), 4.
- Guerrini, S. C. (2006). The developing singer: Comparing the singing accuracy of elementary students on three selected vocal tasks. *Bulletin of the Council for Research in Music Education*, 21-31.
- Gürgen, E. T. (2007). *Orff-Schulwerk ve Kodaly yöntemi'nin vokal doğaçlama, müziksel işitme ve şarkı söyleme becerileri üzerindeki etkileri* [Yayımlanmamış Doktora tezi]. Dokuz Eylül Üniversitesi.
-

- Hale, M. R. (1976). *An experimental study of the comparative effectiveness of harmonic and melodic accompaniment in singing as it relates to the development of a sense of tonality* [Unpublished doctoral dissertation]. The State University of New York.
- Hall, C. (2005). Gender and boys' singing in early childhood. *British Journal of Music Education*, 22(1), 5-20. <https://doi.org/10.1017/S0265051704005960>
- Hayes, A. F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1:1, 77-89, DOI: [10.1080/19312450709336664](https://doi.org/10.1080/19312450709336664)
- Hornbach, C. M., & Taggart, C. C. (2005). The relationship between developmental tonal aptitude and singing achievement among kindergarten, first-, second-, and third-grade students. *Journal of Research in Music Education*, 53(4), 322. <https://doi.org/10.2307/3648430>
- Jaffurs, S. E. (2000). *The relationship between singing achievement and tonal music aptitude* [Unpublished master's thesis]. Michigan State University
- Karaönçel, F. (2021). *Sanal ve akustik çalgı eşlikli şarkı öğretimi uygulamalarının müziksel işitme ve şarkı söyleme becerilerine etkisi* [Yayımlanmamış Doktora tezi]. İnönü Üniversitesi.
- Kendal, J. L. (2018). *The effects of with-text and without-text song presentation styles on preschoolers' singing voice use and pitch accuracy* [Unpublished doctoral dissertation]. University of Maryland.
- Kılıç, S. (2014). Etki büyüklüğü. *Journal of Mood Disorders*, 4 (1), 44-6. DOI: [10.5455/jmood.20140228012836](https://doi.org/10.5455/jmood.20140228012836)
- Kim, S. (2000). *The effects of sequential movement activities on first-grade students' solo singing abilities* [Unpublished master's thesis]. The University of Southern Mississippi
- Leighton, G., & Lamont, A. (2006). Exploring children's singing development: Do experiences in early schooling help or hinder? *Music Education Research*, 8(3), 311-330. <https://doi.org/10.1080/14613800600957461>
- Levinowitz, L. M. (1987). *An experimental study of the comparative effects of singing songs with words and without words on children in kindergarten and first grade* [Unpublished doctoral dissertation]. Temple University.
- Lewinowitz, L. M. Barnes, P. Guerrini, S. Clement, M. D'April, P. Morey, M. J. (1998). Measuring singing voice development in the elementary general music classroom. *Journal of Research In Music Education*, 46 (1), 35-47.
- Mang, E. (2006). The effects of age, gender, and language on children's singing competency. *British Journal of Music Education*, 23(2), 161-174. <https://doi.org/10.1017/S0265051706006905>
- Mizener, C. P. (2008). Our singing children: developing singing accuracy. *General Music Today*, 21(3), 18-24. <https://doi.org/10.1177/1048371308317086>
- Nakamura, Y. (2002). Teacher assessment and peer assessment in practice. *Educational Studies*, 44, 203-215.
- Newlin, G. A. (2004). *The effects of part-work instruction on first-grade part-singing acquisition and achievement*. Unpublished doctoral dissertation]. University of Hartford.
- Nichols, B. E. (2017). Constructing singing assessments for the music classroom. *General Music Today*, 30(3), 13-17. <https://doi.org/10.1177/1048371317690864>

- Pereira, A. I., & Rodrigues, H. (2019). The relationship between Portuguese children's use of singing voice and singing accuracy when singing with text and a neutral syllable. *Music Perception*, 36(5), 468-479. <https://doi.org/10.1525/mp.2019.36.5.468>
- Rutkowski, J. (1986). *The effect of restricted song range on kindergarten children's use of singing voice and developmental music aptitude (elementary, preschool)*. [Unpublished doctoral dissertation]. State University of New York
- Rutkowski, J. (1997). The nature of children's singing voices: Characteristics and assessment. *The phenomenon of singing*, 1, 201-209.
- Rutkowski, J., & Snell Miller, M. (2003a). A longitudinal study of elementary children's acquisition of their singing voices. *Update: Applications of Research in Music Education*, 22(1), 5-14. <https://doi.org/10.1177/87551233020220010901>
- Rutkowski, J., & Miller, M. S. (2003b). The effectiveness of frequency of instruction and individual/small-group singing activities on first graders' use of singing voice and developmental music aptitude. *Contributions to Music Education*, 30 (1), 23-38.
- Rutkowski, J., & Miller, M. S. (2003c). The effect of teacher feedback and modeling on first graders' use of singing voice and developmental music aptitude. *Bulletin of the Council for Research in Music Education*, 1-10.
- Rutkowski, J. (2013). Kindergarten children's use of singing voice achievement and developmental music aptitude in the context of informal music instruction. *Perspectives: Journal of the Early Childhood Music & Movement Association*, 8(1), 6-11. [https://doi.org/10.1386/IJMEC\\_0225\\_1](https://doi.org/10.1386/IJMEC_0225_1)
- Rutkowski, J. (2015). The relationship between children's use of singing voice and singing accuracy. *Music Perception*, 32(3), 283-292. <https://doi.org/10.1525/mp.2015.32.3.283>
- Rutkowski, J. (1990). The measurement and evaluation of children's singing voice development. *The Quarterly*, 1(1-2), 81-95.
- Salvador, K. (2010). How can elementary teachers measure singing voice achievement? a critical review of assessments, 1994-2009. *Update: Applications of Research in Music Education*, 29(1), 40-47. <https://doi.org/10.1177/8755123310378454>
- Smale, M. J. (1987). *An investigation of pitch accuracy of four- and five-year-old singers*. [Unpublished doctoral dissertation]. University of Minnesota
- Smith, J. (2006). Every child a singer: techniques for assisting developing singers. *Music Educators Journal*, 93(2), 28. <https://doi.org/10.2307/3878468>
- Steenwyk, L. A. (2004). *An investigation of the use of singing games in measuring singing voice development in young children*. [Unpublished master's thesis]. Michigan State University.
- Tavani, F. (1998). *The relationships among first grade students' tonal aptitude, singing vocal development, and developmental age*. [Unpublished master's thesis]. Rowan University.
- Tu, C. M. (2009). Music aptitude, tonal accuracy achievement, and singing voice development of kindergarten and first-grade children. *Research Perspectives in Music Education*, 13(1), 29-37.



- Uçal, E. (2003). Okulöncesi müzik eğitiminde Orff öğretisinin müziksel beceriler üzerindeki etkileri [Yayımlanmamış Yüksek Lisans tezi]. Dokuz Eylül Üniversitesi.
- Uysal, G. (2004). *İlköğretimde işbirlikli öğrenmenin müzik öğretiminde sınıf atmosferi ve şarkı söyleme becerileri üzerindeki etkisi* [Yayımlanmamış Yüksek Lisans tezi]. Dokuz Eylül Üniversitesi.
- Vande Wege, R. M. (2005). *The effect of tonal pattern instruction on the singing voice development of first grade students*. [Unpublished master's thesis]. Michigan State University.
- Welch, G. F. (1979). Poor pitch singing: A review of the literature. *Psychology of Music*, 7(1), 50-58. <https://doi.org/10.1177/030573567971006>
- Welch, G. F. (1985). A schema theory of how children learn to sing in tune. *Psychology of Music*, 13(1), 3-18. <https://doi.org/10.1177/0305735685131001>
- Welch, G. F. (1986). A developmental view of children's singing. *British Journal of Music Education*, 3(3), 295-303. <https://doi.org/10.1017/S0265051700000802>
- Welch, G. F., Sergeant, D. C., & White, P. J. (1995). The singing competencies of five-year-old developing singers. *Bulletin of the Council for Research in Music Education*, 155-162.
- Welch, G. F., Sergeant, D. C., & White, P. J. (1997). Age, sex, and vocal task as factors in singing" in tune" during the first years of schooling. *Bulletin of the Council for Research in Music Education*, 153-160.
- Welch, G. F. (2002). Early childhood musical development. L. Bresler & C. M. Thompson (Ed.), *The arts in children's lives* (ss. 113-128). Springer Netherlands. [https://doi.org/10.1007/0-306-47511-1\\_10](https://doi.org/10.1007/0-306-47511-1_10)
- Welch, G. F. (2016). Singing and vocal development. G. McPherson (Ed.), *The child as musician*, (pp.311-329). Oxford University Press.
- Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance, and self-assessment. *Musicae Scientiae*, 12(1), 3-26. <https://doi.org/10.1177/102986490801200102>

## **ŞARKI SÖYLEME TONUNUN GELİŞİMİ ÖLÇEĞİNİN TÜRKÇEYE UYARLANMASI ve ÇOCUKLARIN ŞARKI SÖYLEME TONU KULLANIMININ FARKLI DEĞİŞKENLER AÇISINDAN İNCELENMESİ**

### **ÖZ**

Bu araştırmada Rutkowski tarafından geliştirilmiş “Şarkı Söyleme Tonunun Gelişimi Ölçeği”nin Türkçe ’ye uyarlanması ve çocukların şarkı söyleme tonun yaş, cinsiyet ve sözlü/sözsüz müziklerin kullanımı gibi farklı değişkenler açısından incelenmesi amaçlanmıştır. Bu amaç doğrultusunda gerekli izinler alındıktan sonra ölçme aracının çevirisi bağımsız iki çevirmen tarafından yapılmış ve dilsel eşdeğerlik çalışması kapsamında uzman görüşlerine sunularak çeviri son halini almıştır. Araştırma grubunu 2022/2023 eğitim ve öğretim yılında Ankara’nın Çankaya ilçesindeki bir ilkokulda anasınıfı ve 3. sınıf arasında eğitimine devam eden 253 çocuk oluşturmuştur. Rasch modeline göre yapılan analizlerin sonucunda ölçme aracının amaç doğrultusunda çalıştığı, geçerliğine ilişkin olumlu bulgular içerdiği ve amaç doğrultusunda kullanılabilceği ortaya koyulmuştur. Ölçme aracının güvenilirliğine ilişkin ise puanlayıcılar arası ve puanlayıcı içi güvenilirlik değerlerinin oldukça yüksek olduğu belirlenmiştir. Bu bulgular Şarkı Söyleme Tonunun Gelişimi Ölçeği’nin Türk kültüründe geçerli ve güvenilir bir ölçme aracı olduğunu göstermektedir. Türk çocuklarının şarkı söyleme tonunu kullanım durumları farklı değişkenler açısından incelendiğinde ise sınıf düzeyi arttıkça şarkı söyleme tonu kullanım düzeyinin arttığı görülmektedir. Ayrıca kız çocuklarının erkek çocuklarına kıyasla şarkı söyleme tonu kullanımında nitekim ileri olduğu ve sözsüz şarkıların, sözlü şarkılara kıyasla çocukların şarkı söyleme tonu kullanımında avantaj sağladığı belirlenmiştir. Şarkı Söyleme Tonunun Gelişimi Ölçeği’nin genel müzik eğitimi sınıflarında kullanımına yönelik olarak eğitim sistemimize dâhil edilmesi önerilmektedir. Yanı sıra eğitim sistemimizdeki mevcut eksikliklerin giderilmesi ve iyileştirilmesi amacıyla çocukların şarkı söyleme tonunun gelişimini farklı değişkenler açısından ölçmeye yönelik, farklı sınıf düzeylerinde geniş örneklemli tarama çalışmalarının yapılması önerilmektedir.

**Anahtar kelimeler:** Şarkı Söyleme Tonunun Gelişimi Ölçeği, genel müzik eğitimi, şarkı söyleme, geçerlik, güvenilirlik.

## GİRİŞ

Genel müzik eğitimi, temel müzikal becerileri, bilgileri ve anlayışı geliştirmek amacıyla ayrıt etmeksizin tüm öğrencilere yönelik tasarlanmış müzik eğitiminin bir türünü ifade etmektedir. İlköğretim ve ortaöğretim müfredatlarının bir parçası olan genel müzik eğitimi günümüzde dünya çapında hem müzik eğitimcileri hem de sınıf öğretmenleri tarafından yürütülmektedir. Kapsamında söyleme, çalma, yaratıcılık ve müzik ile etkileşim kurmaya yönelik öğrenme hedefleri yer almaktadır (Abril, 2016, s. 5). Şüphesiz ki bu öğrenme hedefleri arasında genel müzik eğitimi sınıflarının en temel etkinliği şarkı söylemedir. Müzik eğitimcisinin en önemli görevi genel müzik eğitiminin başladığı ilk yıllardan itibaren çocuklara şarkı söylemeyi öğretmek ve çocukların şarkı söyleme isteklerini teşvik etmektir (Mizener, 2008, s. 23). Bu amaçla müzik eğitimcisinin çocuğun şarkı söyleme tonunun gelişimsel doğasını anlaması ve bireysel ihtiyaçlara göre bir öğretim programı geliştirmesi önemlidir.

Şarkı söyleme becerilerinin temeli anne karnında atılmaktadır. Annenin konuşma ve şarkı söyleme sesi, ses dalgaları aracılığıyla fetüse aktarılmaktadır. Gelişimi sürmekte olan fetüsün gebeliğin son aylarındaki işitsel ve duygusal deneyimleri müzikal etkileşimin ve şarkı söyleme becerilerinin temel kaynağını oluşturmaktadır (Welch, 2016, s. 442). Doğumdan önce başlayan müzikal gelişim, doğum sonrasında ise bebeklere/çocuklara sunulan müzikal çevrenin niteliği doğrultusunda şekillenmektedir. Bebekler hem konuşma hem de şarkı söyleme tonunu kullanma potansiyeli ile doğmaktadır. Konuşma tonu sürekli olarak yetişkinler tarafından modellenmekte ve bu sayede bebekler konuşma becerilerini hızla geliştirebilmektedir. Ne yazık ki şarkı söyleme becerileri, konuşma becerilerinin gelişimin gerisinde kalmaktadır. Bunun temel sebebi anadilin konuşulduğu çevrede bebeklerin şarkı söyleme tonuna yeterince maruz kalmamasıdır. Böylece bebekler şarkı söyleme tonunu keşfetmek yerine konuşma tonunu kullanmaya daha yatkınlık göstermektedir (Gordon, 2007, s. 18-19). Bu bağlamda doğum itibarıyla bebeklere, anadilini öğrendikleri enformel yollarla müzik dilini de öğrenebilecekleri bir müzikal çevrenin sunulması büyük önem taşımaktadır. Anadilinde belirli bir kelime haznesine sahip çocukların konuşma becerisini elde ettiği gibi, belirli bir müzikal kelime haznesine sahip çocuklar da çok küçük yaşlardan itibaren şarkı söyleme tonunu elde edebilmektedir. Gordon (2013, s. 142)'a göre konuşma tonu nasıl çevredeki modelleri dinleyerek ve taklit ederek öğreniliyorsa; şarkı söyleme tonu da benzer yollarla iyi bir modeli canlı etkileşimlere dayalı olarak dinleyerek ve taklit ederek öğrenilmektedir.

İdeal olarak genel müzik eğitime başlamadan önce 0-6 yaş aralığında zengin bir müzikal kelime haznesinin oluşturulmuş olması ve odiasyon becerilerinin gelişmiş olması gerekmektedir; ancak her çocuğun erken yaşlardan itibaren enformel bir müzik öğrenimine erişimi mümkün olmamaktadır. Bu durumda anasını ve ilkokula başlamış olan çocukların şarkı söyleme tonunu keşfedemediği veya zayıf bir entonasyona sahip olduğu görülmektedir (Gordon, 2007, s. 51). Gerekli müzikal alt yapısı olmadan genel müzik eğitime başlayan çocuklar şarkı söyleme alanında çeşitli sorunlarla karşılaşabilmektedir. Welch (1986, s. 300) genel müzik eğitimi sınıflarında çocukların şarkı söyleme hususunda karşılaşabildikleri sorunları şu şekilde sıralamaktadır:

1. Çocuklar grup içinde kendi seslerini ve tepkilerini duymakta zorluk çekebilmektedir,

2. Bireysel olarak şarkı söyleme ve bireysel olarak yardım alma imkânının az olmasından kaynaklı zorluklar yaşanabilmektedir,
3. Çocuklar okul müziği repertuarına başlamadan sınırlı bir ses bölgesinde çalışmaya ve o ses bölgesi içinde güven kazanmaya ihtiyaç duymaktadır. Bu pratiğin yokluğundan kaynaklı zorluklar yaşanabilmektedir.
4. Çocuklar müzik ve sözleri birbirinden ayrı çalışmaya ihtiyaç duyabilmektedir (Özellikle bazı çocuklar için dilin karmaşıklığı olmadan doğal hecelerle sadece müzik üzerinde çalışılması ihtiyacı bulunmaktadır).
5. Ve çocuklar nefes kontrolü gibi hususlar üzerinde pratik yaparak sesini daha etkili nasıl kullanabileceğini öğrenmeye ihtiyaç duymaktadır.

Geçmişten bu yana şarkı söyleme tonunu kullanamayan bireyler literatürde çok çeşitli sıfatlarla tanımlanmaktadır. Bunlar arasında monoton, belirsiz-şarkıcı, akortsuz şarkıcı, zayıf perdeli şarkıcı, problem şarkıcı, kusurlu şarkıcı gibi sıfatlar yer almaktadır (Aaron, 1990, s. 4). Bu sıfatların çoğu şarkı söyleme becerilerinin gelişimsel özelliklerini vurgulamamaktadır. Aslında şarkı söyleme tonunun kullanılmaması doğuştan gelen, sesleri ayırt etme becerisinin yoksunluğu ile ilişkili bir engel değil, çeşitli çevresel faktörlerin ürünüdür (Welch, 1979, s. 56). Genel kanı şarkı söylemenin öğretilebilir ve öğrenilebilir bir beceri olduğudur (Welch, 1985, 1986; Gould, 1968, 1969; Smith, 2006; Mizener, 2008; Rutkowski & Miller, 2003c). Welch (1985, s. 3) şarkı söyleme becerilerinin gelişimsel doğasını şu sözleri ile ifade etmektedir; “Çocuklar için, ton içinde şarkı söylemek sadece bir "yapabilir/yapamaz" meselesi değil, daha ziyade çok çeşitli yeterlilik derecelerine sahip bir hiyerarşi veya beceri sürekliliğidir.” Bu bağlamda şarkı söyleme tonunu kullanamayan çocuklar (akortsuz şarkıcılar) eğitimden yararlanabilen, gelişmekte olan şarkıcılar olarak değerlendirilmektedir (Leighton&Lamont, 2006, s. 311).

Gould (1968, s. 8) ise küçük çocukların şarkı söylemesi önündeki en büyük engelin sosyal çevrenin sebep olduğu psikolojik engellerden kaynaklandığını şu sözleri ile vurgulamaktadır;

“Aslında okul öncesi ve anasınıfı çağında şarkı söylemeyen bir çocuğun var olup olmadığı şüphelidir. Çocukların hepsi şarkı söyleyebilmektedir. Belki bazıları öğretmenin istediği şekilde söyleyememektedir; ama hepsi şarkı söyleyebilmektedir. Problemlili şarkıcılar, şarkı söyleme çabalarına karşı akranlarından, öğretmenlerinden veya ebeveynlerinden olumsuz tepkiler aldıktan sonra psikolojik engellemeler gelişmektedir.”

Müzik öğrenimi sürecinde erken yaşlardan itibaren şarkı söyleme tonunu keşfedemeyen veya kullanmayı öğrenmeyen çocuklar büyüdükçe müzikten uzaklaşmakta ve müziğe karşı yabancılaşmaktadır (Gould, 1969, s. 11). Çocukların şarkı söyleme becerilerinin hangi evrede olduğunun tespit edilmesi, bireysel ihtiyaçlarının belirlenmesi ve bu doğrultuda iyileştirici tekniklerin ve telafi edici bir rehberliğin uygulanması kritik önem taşımaktadır. Bu noktada amaca yönelik doğru ölçme aracının kullanılması gerekmektedir. Salvador (2010, s. 43) şarkı söyleme davranışını ölçmeye yönelik bir araç tasarlarırken veya uyarlama yaparken ölçme aracının şarkı söyleme kazanımlarını ölçmeye yönelik geçerli bir araç olması, ölçme aracının kabul edilebilir düzeyde güvenilirlik değerleri sağlıyor olması, ölçütlerin açık ve anlaşılır olması, ölçme aracının sınıf ortamına dâhil edilebilecek

nitelikte olması ve her bir çocuğun ihtiyaçlarının belirlenmesine yönelik tanı koyabilecek nitelikte olması gibi hususlara dikkat edilmesi gerektiğini vurgulamaktadır.

Uluslararası literatür incelendiğinde şarkı söyleme davranışına yönelik ölçme araçlarının, şarkı söyleme tonu gelişimini ölçmeye yönelik araçlar (Rutkowski, 1990, 1997) ve şarkı söyleme davranışının doğruluğunu (entonasyona yönelik) ölçmeye yönelik araçlar (Boardman, 1964; Dittmore, 1968; Hale, 1976; Newlin, 2003; Hornbach&Taggart, 2005; Wise&Sloboda, 2008) olmak üzere ikiye ayrıldığı görülmektedir. Nichols (2017, s. 14) da şarkı söylemeye yönelik araçların ölçülmek istenen özelliğe yönelik tasarlanması ve kullanılması gerektiğini vurgulamaktadır. Eğer amaç öğrencilerin şarkı söyleme tonu gelişimini ölçmek ise ses aralığının ve şarkı söyleme tonu veya konuşma tonu kullanımı gibi şarkı söyleme davranışının gelişimsel özelliklerini ölçen araçlardan; amaç belirli perdelerin iyi bir entonasyonla/ton içinde seslendirilmesinin ölçülmesi ise şarkı söyleme davranışının doğruluğunu ölçen araçlardan yararlanılmalıdır. Şarkı söyleme tonunun gelişimi ve vokal doğruluk birbirinden ayrı; fakat birbiri ile bağlantılı becerilerdir. Rutkowski (2015) anasınıfı ve birinci sınıf çocuklarının şarkı söyleme tonu kullanım durumu ile vokal doğrulukları arasında pozitif yönde bir korelasyon olduğunu ortaya koymuştur. Bu çalışmadan elde edilen sonuçları henüz vokal ses bölgesini kullanamayan çocukların ne kadar iyi şarkı söyleyebildiğine odaklanılmasının (vokal doğruluk, ton içinde şarkı söyleme) şarkı söyleme becerilerinin gelişimi açısından faydalı olmayacağını öne sürmektedir. Bu doğrultuda çocukların vokal doğruluğuna ilişkin ölçümler yapılmadan önce, şarkı söyleme tonu (vokal ses bölgesinin) kullanım durumlarının belirlenmesi önemlidir.

Şarkı söyleme tonunun kazanımını evrelere ayırarak gelişimsel modeller sunan iki önemli çalışma bulunmaktadır (Welch, 2002; Rutkowski, 1997). Welch (2002) dört evreden oluşan “vokal ses-eşleştirme gelişimi” modelini geliştirmiştir. Bu modele göre birinci evrede melodiden ziyade şarkının sözleri ön plandadır, melodi çoğunlukla ritmik şarkı söyler gibi seslendirilmektedir ve kısıtlı bir ses aralığı kullanılmaktadır. İkinci evrede ses perdesindeki değişimlerin kontrol edilebileceğine yönelik farkındalık artmakta, melodik çizgi genel hatlarıyla takip edilebilmekte ve kullanılan ses aralığı genişlemektedir. Üçüncü evrede melodik doku ve aralıklar çoğunlukla doğruluk kazanmakta; ancak tonalitede kimi değişimler meydana gelebilmektedir. Dördüncü evrede ise melodik dokuda veya seslerde kayda değer bir hata olmamakta, şarkı söyleme tonu daha kesinlik kazanmaktadır.

Rutkowski ise on beş yılı aşkın bir süre boyunca çocukların şarkı söyleme davranışını sistematik olarak değerlendirmiş ve dokuz aşamalı bir model sunmuştur (Welch, 2016, s. 446). Bu modele göre çocuklar konuşma ses bölgesinin kullanımından, sınırlı bir ses bölgesinin, başlangıç ses bölgesinin ve genişletilmiş ses bölgesinin kullanımına doğru dokuz aşamalı bir ilerleme göstermektedir. Şarkı Söyleme Tonunun Gelişimi Ölçeği de bu model üzerine kurulmuştur. Şarkı Söyleme Tonunun Gelişimi Ölçeği genel müzik eğitimi sınıflarında kullanıma uygun, pratik, anlaşılır ve güvenilirliği sağlam bir şekilde ortaya konmuş bir ölçme aracıdır. Ses aralığı re3-si bemol 3 ve üzerindeki aralıklardan oluşan şarkıların performansının değerlendirilmesinde kullanılabilir. Bunun yanı sıra şarkı söyleme tonunu gelişimsel bir perspektifte değerlendirmekte ve çocukların kazanımlarını arttırmaya yönelik nelere ihtiyacı olduğuna ilişkin teşhis edici bilgi sunmaktadır (Salvador, 2010, s. 44-45)

Ulusal literatür tarandığında şarkı söyleme becerilerini ölçmeye yönelik tasarlanmış ölçme araçlarının şarkı söyleme davranışının gelişimsel özelliklerinden ziyade çoğunlukla şarkı söylemede doğruluğa ve çeşitli açılardan şarkı söyleme davranışlarının gelişimine odaklandığı görülmektedir (Bilen, 1995; Uçal, 2003; Gürgen, 2007; Uysal, 2004; Karaönçel, 2021). Bu araçlar arasında yalnızca Değer (2012) öğrencilerin şarkı söyleme tonunun gelişimini ölçmeye yönelik “Şarkı Sesinin Kullanımında Davranış Gelişim Ölçeği”ni geliştirmiş, geçerlik ve güvenilirliğini ortaya koymuştur. Bu ölçek konuşma tonunun kullanımından, altılı, sekizli ve üzeri aralıklardan oluşan çeşitli ses bölgelerinin kullanımına kadar şarkı söyleme tonunun gelişimini belirlemeye yönelik tasarlanmıştır. Literatüre bakıldığında geçerlik ve güvenilirliği geniş bir örnekleme ve detaylı bir şekilde ele alınmış, şarkı söyleme tonunu gelişimsel perspektifte değerlendiren bir ölçme aracına ihtiyaç duyulduğu belirlenmiştir. Rutkowski’nin “Şarkı Söyleme Tonunun Gelişimi Ölçeği” geçerliği ve güvenilirliği kesin olarak ortaya koyulmuş, uluslararası literatürde çok geniş kullanım oranına sahip bir ölçme aracıdır (Tavani, 1998; Jaffurs, 2000; Kim, 2000; Guerrini, 2002, 2004, 2006; Doneski, 2005; Mang, 2006; Vande Wege, 2005, Steenwyk, 2004; Hall, 2005; Tu, 2009; Kendal, 2018; Pereira & Rodrigues, 2019). Buradan yola çıkarak mevcut araştırmada Şarkı Söyleme Tonunun Gelişimi Ölçeği’nin Türkçe’ye uyarlanması, performans dayalı bir ölçme aracı olarak geçerlik ve güvenilirliğinin sınanması ve şarkı söyleme tonu gelişiminin çeşitli değişkenler açısından değerlendirilmesi amaçlanmıştır. Bu amaç doğrultusunda aşağıdaki sorular sırasıyla ele alınmıştır:

1. Şarkı Söyleme Tonunun Gelişimi Ölçeği sözlü ve sözsüz değerlendirme ölçütlerine ilişkin geçerlik ve güvenilirlik değerleri nasıldır?
2. Öğrencilerin sözlü ve sözsüz değerlendirme ölçütlerine ait performansları, buldukları sınıf düzeylerine göre anlamlı bir farklılık göstermekte midir?
3. Öğrencilerin sözlü ve sözsüz değerlendirme ölçütlerine ait performansları cinsiyetlerine göre anlamlı bir farklılık göstermekte midir?
4. Öğrencilerin sözlü ve sözsüz değerlendirme ölçütlerine ilişkin performansları arasında istatistiksel olarak anlamlı bir ilişki var mıdır?

## YÖNTEM

### Araştırma Grubu

Araştırma grubunu 2022/2023 eğitim-öğretim yılında Ankara'nın Çankaya ilçesindeki bir ilkokulda anasınıfları ile 3. sınıf arasında eğitimine devam eden 253 çocuk oluşturmuştur. Araştırmaya katılımda gönüllülük esası temel alınmış, Gazi Üniversitesi Etik Komisyonu'nun (10.03.2023 tarih ve 2023 – 248 sayılı karar) öngördüğü onam formları uygulama öncesi ebeveynlerden toplanmıştır. Katılımcıların cinsiyet ve sınıf düzeylerine göre dağılımı aşağıdaki tabloda sunulmuştur.

**Tablo 1:** Katılımcıların Cinsiyet ve Sınıf Düzeyine Göre Dağılımı.

Sınıf	Erkek	Kız	Toplam
Anasınıfları	7	11	18
1. Sınıf	55	35	90
2. Sınıf	42	48	90
3. Sınıf	26	29	55
Toplam	130	123	253

### Veri Toplama Aracı

#### Şarkı Söyleme Tonunun Gelişimi Ölçeği (SVDM)

Bu ölçek çocukların şarkı söyleme tonunu kullanım durumunu ölçmek amacıyla Rutkowski (1990, 1997) tarafından uzun yıllar süren geliştirme ve revizyon çalışmaları sonucunda literatüre kazandırılmıştır. Sorular re-si bemol aralığında sekiz adet motiften oluşturulmuştur. Uygulama öncesinde ön hazırlık olarak sınıfça belirlenmiş olan bu motifler üzerinde bir süre çalışılmaktadır. Ön hazırlık ardından sorular her bir çocuğa bireysel olarak ayrı bir sınıf ortamında uygulanmaktadır. Bu sınıfa tonu ve motifleri duyurabilmek için gerekli olan piyano/org/ton çanı gibi bir çalgı ve uzmanların değerlendirme yapabilmesi için bir kayıt cihazı yerleştirilmelidir. Uygulama prosedürü sırasıyla şu basamaklarla gerçekleşmektedir;

1. İlk motif piyanoda çalınır (örneğin: la-fa-re).
2. Ardından aynı motif öğretmen tarafından natürel heceler (bum, pam v.b.) veya sözleri ile seslendirilir. Bu aşamada eşlik kullanılmaz.
3. Çocuktan motifi yansılması istenir. Bu aşamada da herhangi şekilde eşlik edilmez.
4. İlk 3 aşama her bir motif için tekrar edilir.
5. Basamaklar arasında bekleme süresi bırakılmaz.
6. Grubun bir yarısı motifleri önce natürel hecelerle, sonra sözler ile seslendirir; diğer yarısı ise önce sözlerle ardından natürel hecelerle seslendirir (Rutkowski ve Miller, 2003a, s.6-7).



Şekil 1: Şarkı Söyleme Tonunun Gelişimi Ölçeği Motifleri.

Tablo 2: Derecelendirme Ölçeği (Rutkowski & Miller, 2003a, s. 8)

1	"Ön-Şarkıcı", şarkıyı (motifi) söyleyemez, yalnızca ritmik olarak sözleri seslendirir.
1,5	"Tutarsız, konuşma ses bölgesinde şarkı söyleyen", bazen ritmik olarak şarkıyı söyler, bazen sesleri verir ve ses yüksekliğine karşı bir miktar hassasiyet sergiler; ancak konuşma ses bölgesinde kalır (genellikle la2-do3 aralığında).
2	"Konuşma ses bölgesinde şarkı söyleyen" sesleri verir ve ses yüksekliğine karşı bir miktar hassasiyet gösterir, ancak konuşma ses bölgesinde kalır (genellikle la2-do3 aralığında).
2,5	"Tutarsız, sınırlı bir ses bölgesinde şarkı söyleyen" konuşma ve şarkı söyleme tonu arasında gidip gelir ve şarkı söyleme tonundayken sınırlı bir aralık kullanır (genellikle Fa3'e kadar).
3	"Sınırlı ses bölgesinde şarkı söyleyen", sınırlı ses bölgesinde şarkı söyleme tonunu tutarlı bir şekilde kullanır (genellikle re3 - fa3 aralığında).
3,5	"Tutarsız, başlangıç ses bölgesinde şarkı söyleyen", kimi zaman sınırlı ses bölgesinde şarkıyı seslendirirken, kimi zaman başlangıç ses bölgesini kullanır (genellikle re3-la3 aralığını kullanır).

4	"Başlangıç ses bölgesinde şarkı söyleyen" başlangıç ses bölgesini tutarlı bir şekilde kullanır (genellikle re3-la3 aralığını kullanır).
4,5	"Tutarsız Şarkıcı" zaman zaman yalnızca başlangıç ses bölgesini kullanır, zaman zaman ise genişletilmiş ses bölgesini kullanır (si bemol 3 ve üzeri).
5	"Şarkıcı", tutarlı bir şekilde genişletilmiş ses bölgesini kullanır (si bemol 3 ve üzeri).

### Verilerin Analizi

Bu araştırmanın veri analizi iki aşamadan oluşmaktadır. Birinci aşamada; ölçme aracının geçerliliğini test etmek amacıyla çok yüzeysel Rasch ölçme modeli hesaplanmıştır. Araştırmada değerlendirme (sözlü ve sözsüz), puanlayıcılar, öğrenci, sınıf, cinsiyet olmak üzere beş yüzeysel bulunmaktadır.

İkinci aşamada; öğrencilerin sözlü ve sözsüz performanslarına ait puan ortalamalarının sınıf ve cinsiyet değişkenlerine göre farklılaşıp farklılaşmadığı tek yönlü varyans analizi ve ilişkisiz t testi ile belirlenmiştir. Öğrencilerin sözlü ve sözsüz ölçütlere ait performansları arasında anlamlı bir ilişkinin olup olmadığı ise Pearson momentler çarpımı korelasyon katsayısı ile belirlenmiştir.

### Uyarlama Süreci: Ölçek Maddelerinin Türkçeye Çevrilmesi

Joenne Rutkowski'den uyarlama izni alındıktan sonra ölçeğin İngilizceden Türkçeye çeviri süreci başlamıştır. Orijinal form, her iki dile de hâkim, bağımsız bir yeminli tercüman ve araştırmacı tarafından Türkçeye çevrilmiştir. İlk çeviriden elde edilen ifadeler birbirleriyle karşılaştırılarak ortak ifadeler tespit edilmiştir. Maddelerin orijinal dilinde ve iki farklı çeviri sonucu elde edilen ortak ifadelerin yer aldığı uzman görüş formu, Gazi Üniversitesi, Gazi Eğitim Fakültesi, Müzik Eğitimi Anabilim Dalı ve İngilizce Eğitimi Anabilim Dalı'nda görev yapan 6 öğretim üyesinin görüşlerine sunulmuştur. Dilsel eşdeğerlik çalışması kapsamında uzmanlardan alınan görüşler doğrultusunda Türkçe maddeler üzerinde gerekli düzenlemeler yapılmış ve İngilizce'den Türkçeye çeviri son halini almıştır.

## BULGULAR

### Şarkı Söyleme Tonunun Gelişimi Ölçeği Sözlü ve Sözsüz Ölçütlerine İlişkin Geçerlik ve Güvenirlik Değerleri Nasıldır? Sorusuna İlişkin Bulgular ve Yorumlar

Bu araştırma sorusunda geçerlik kanıtı olarak her iki ölçüte ilişkin ölçüm değerlerine bakılmıştır. Öğrencilerin sözlü ve sözsüz ölçütlere ilişkin performans ölçümleri çok yüzeysel Rasch ölçme modeline göre analiz edilmiştir. Ölçütlere ait değerler Tablo 3'te verilmiştir.

**Tablo 3.** Sözlü ve Sözsüz Ölçüt Yüzeyine Ait Ölçüm Raporu

Ölçüt	Gözlenen ortalama	Düzeltilmiş Ortalama	Logit Değeri	Standart Hata	Uyum İçi	Uyum Dışı
Sözlü	3,196	3,243	-0,23	0,07	0,98	0,98
Sözsüz	3,309	3,424	0,23	0,07	1,00	0,97
Ortalama	3,252	3,333	0,00	0,07	0,99	0,98



S (Evren)	0,056	0,090	0,23	0,00	0,01	0,01
S (Örneklem)	0,080	0,128	0,33	0,00	0,01	0,01
Model, Evren		RMSE= 0,07	Düzeltilmiş S.H.= 0,22	Ayırma Oranı= 3,13		
			Ayırma İndeksi= 4,51	Güvenirlik= 0,91		
Model, Örneklem		RMSE= 0,07	Düzeltilmiş S.H.= 0,32	Ayırma Oranı= 4,54		
			Ayırma İndeksi= 6,39	Güvenirlik= 0,95		
Model, Ki-kare (Sabit etkili) :			21,6	sd= 1	p= 0,00	
S: Standart sapma, RMSE: Hata kareleri ortalamasının karekökü, S.H.: Standart hata, sd: Serbestlik derecesi						

Tablo 3 incelendiğinde, sözlü ve sözsüz ölçütlerine ait logit değerlerinin -0,23 ve 0,23 olduğu ve ölçütlerin 0,46 [0,23-(-0,23)] logit değişime sahip olduğu görülmektedir. Ölçütlere ait logit değerleri ortalaması 0,00 ve standart sapması 0,32'dir. Bu değerler sözsüz değerlendirmenin gruba göre kısmen kolay, sözlü değerlendirmenin ise kısmen zor olduğunu işaret etmektedir. Kriter yüzeyine ait verilere bakıldığında ise uyum içi 0,98 ve 1,00 iken uyum dışı 0,98 ve 0,97'dir. Uyum içi değerlerin ortalaması 0,99 ve uyum dışı değerlerin ortalaması 0,98'dir. Uyum içi ve uyum dışı değerlerin 1'e yakın olması model veri uyumunun iyi olduğunu göstermektedir. Kriter yüzeyine ait ayırma indeksinin 6,39 olduğu görülmektedir. Ayırma indeksinin 2,00'nin üzerinde olması ölçütlerin farklı seviyelerdeki bireylere uygulanabileceğini ve ölçütlerin işlevsel olduğunu göstermektedir (Nakamura, 2002). Ayırma oranı ve güvenilirlik indeksi değerlerinin yüksek olması ölçütlerin güçlük düzeylerinin farklı olduğunu göstermektedir. Ki kare istatistiği incelendiğinde bu farkın istatistiksel açıdan anlamlı olduğu görülmektedir ( $\chi^2(sd) = 21,6 (1), p = 0,00 < 0,01$ ). Bu durum ölçüt güçlükleri arasında anlamlı fark olduğunu göstermektedir. Sözlü ve sözsüz ölçütlerin logit değeri, uyum içi ve uyum dışı değerleri ve ayırma indeksi aracın amaç doğrultusunda çalıştığını, geçerliğine ilişkin olumlu bulgular içerdiğini ve amaç doğrultusunda kullanılabileceğini göstermektedir.

Sözlü ve sözsüz ölçütlere yönelik güvenilirlik kanıtı olarak puanlayıcılar arası güvenilirliğe bakılmıştır. Puanlayıcılar arası güvenilirlik Krippendorff alfa katsayısı kullanılarak hesaplanmıştır (Hayes & Krippendorff, 2007). Bulgulara göre sözlü kategori için güvenilirlik katsayısı  $\alpha = .99$ , sözsüz kategori için  $\alpha = .98$  olarak elde edilmiştir. Bu bulgular puanlayıcılar arası güvenilirliğin oldukça yüksek düzeyde olduğunu göstermektedir. Puanlayıcı içi güvenilirlik değerleri ise Pearson Korelasyon Katsayısı ile belirlenmiştir. Buna göre birinci puanlayıcının iki farklı zamanda gerçekleştirdiği değerlendirmeleri arasındaki korelasyon katsayısı sözlü kategori için  $r = .94$ , sözsüz kategori için  $r = .98$  olarak; ikinci puanlayıcının değerlendirmeleri arasındaki korelasyon katsayısı ise sözlü kategori için  $r = .96$ , sözsüz kategori için  $r = .99$  olarak belirlenmiştir. Bu değerler puanlayıcı-ici güvenilirliğin de oldukça yüksek düzeyde olduğunu göstermektedir.

### **Öğrencilerin Sözlü ve Sözsüz Ölçütlere Ait Performansları, Buldukları Sınıf Düzeylerine Göre Anlamlı Bir Farklılık Göstermekte Midir? Sorusuna İlişkin Bulgular ve Yorumlar**

Öğrencilerin sözlü ve sözsüz ölçütlere ait performans puanlarının buldukları sınıf düzeyine göre anlamlı bir farklılık gösterip göstermediğine; test değişkenine ait puanların sınıf düzeylerine normal dağılım göstermesi, puanların eşit aralıkta olması ve gruplara ait varyans değerlerinin homojen olması nedeniyle tek yönlü varyans

analizi ile bakılmıştır. Analizler 253 çocuk üzerinden yapılmıştır. Analiz sonucu elde edilen bulgular aşağıda Tablo 4'te verilmiştir.

**Tablo 4.** Öğrencilerin Sözlü ve Sözsüz Ölçütlere Ait Performanslarının Sınıf Düzeylerine Göre Tek Faktörlü Varyans Analizi İle Karşılaştırılması

Boyut	Kategori	N	Ort	SS	F	P	Fark (Tukey)	Eta kare
sözlü	anasınıfı	18	2,35	1,05	6,047	,001	Anasınıf-2.S Anasınıf-3.S	0,28
	1.sınıf	90	3,03	1,19				
	2.sınıf	90	3,45	1,27				
	3.sınıf	55	3,53	1,20				
	<b>Toplam</b>		253	3,24	1,25			
sözsüz	anasınıfı	18	2,44	1,20	4,981	,002	Anasınıf-2.S Anasınıf-3.S	0.26
	1.sınıf	90	3,18	1,22				
	2.sınıf	90	3,58	1,30				
	3.sınıf	55	3,52	1,29				
	<b>Toplam</b>		253	3,34	1,29			

Farklı sınıf düzeylerinde eğitim gören öğrencilerin sözlü ölçüte ait performansları sınıf düzeylerine göre karşılaştırılmış ve gruplar arası farklılıkların istatistiksel olarak anlamlı olduğu görülmüştür ( $F(3, 249) = 6.047, p < .05$ ). Diğer bir ifade ile sözlü değerlendirmenin sınıf düzeylerine göre farklılık gösterdiği belirlenmiştir. Farklılıkların hangi ikili gruptan kaynaklandığını gösteren çoklu karşılaştırma Tukey (HSD) testi sonuçları incelendiğinde ana sınıfı (Ort.= 2.35; SS=1.05) ve birinci sınıf (Ort.= 3.03; SS=1.19) öğrencilerinin sözlü değerlendirme puanları arasında görece farklılık bulunmuştur; ancak bu fark istatistiksel olarak anlamlı değildir. Anasınıfı ile ikinci sınıf (Ort.= 3.45; SS=1.27) ve anasınıfı ile üçüncü sınıf (Ort.= 3.53; SS=1.20) öğrencilerinin sözlü ölçüte ilişkin performansları arasında ise anlamlı bir farklılık olduğu tespit edilmiştir ( $p < .05$ ).

Öğrencilerin sözsüz ölçüte ilişkin performansları sınıf düzeylerine göre karşılaştırıldığında; gruplar arasında istatistiksel olarak anlamlı fark bulunmuştur ( $F(3, 249) = 4.981, p < .05$ ). Diğer bir ifade ile sözsüz ölçüte ilişkin performanslar sınıf düzeyine göre değişmektedir. Farklılıkların hangi ikili gruptan kaynaklandığını belirlemek amacıyla çoklu karşılaştırma Tukey (HSD) testi yapılmıştır. Sonuçlar incelendiğinde ana sınıfı (Ort.= 2.44; SS=1.21) ve birinci sınıf (Ort.= 3.18; SS=1.22) öğrencilerinin sözsüz değerlendirmeleri arasında anlamlı bir farklılık bulunmamıştır. İlkokul ikinci sınıf (Ort.= 3.58; SS=1.30) ile üçüncü sınıf (Ort.= 3.52; SS=1.29) öğrencilerinin sözsüz değerlendirme puanları arasında görece bir farklılık olduğu belirlenmiştir ( $p < .05$ ). Sözsüz ölçütte üçüncü sınıf performanslarının görece ikinci sınıf düzeylerine göre düştüğü görülse de bu fark istatistiksel olarak anlamlı değildir.

Sınıf düzeylerine göre sözlü ve sözsüz performans ortalamaları incelendiğinde öğrencilerin sözsüz ölçüte ilişkin performanslarının, sözlü ölçüte ilişkin performanslardan görece daha yüksek olduğu görülmektedir.

Gruplara ait veriler arasındaki farkın kayda değer olup olmadığını gösteren bir başka ölçüt ise etki büyüklüğüdür (Kılıç, 2014). Her iki değerlendirme için hesaplanan etki büyüklük değeri 0.8'den büyük olduğu için kuvvetli etki olarak tanımlanmaktadır (Cohen,1988). Buradan da etki büyüklüklerinin benzer düzeyde oldukları görülmektedir.

### **Öğrencilerin Sözlü ve Sözsüz Ölçütlere Ait Performansları Cinsiyetlerine Göre Anlamli bir Farklilik Göstermekte Midir? Sorusuna İlişkin Bulgular ve Yorum**

Öğrencilerin sözlü ve sözsüz ölçütlere ait performans puanlarının cinsiyetlerine göre anlamli bir farklılık gösterip göstermediğine, gruplara ait puanların normal dağılım göstermesi ve puanların eşit aralıklı düzeyde olması nedeniyle ilişkisiz t-testi ile bakılmıştır. Analizler 253 çocuk üzerinden yapılmıştır. Analiz sonucu elde edilen bulgular Tablo 5'te verilmiştir.

**Tablo 5.** Öğrencilerin Sözlü ve Sözsüz Ölçütlere Ait Performanslarının Cinsiyetlerine Göre İlişkisiz T-Testi İle Karşılaştırılması

	Cinsiyet	N	Ort.	SS	SD	t	p
Sözlü	kız	123	3,53	1,19	251	3,600	,000
	erkek	130	2,97	1,25			
Sözsüz	kız	123	3,62	1,25	251	2,721	,001
	erkek	130	3,08	1,28			

Öğrencilerin sözlü ve sözsüz ölçüte ilişkin performanslarında cinsiyetlerine göre anlamli bir farklılık bulunmuştur ( $t_{(253)} = 3,600; p < 0.5$ ). Kız öğrencilerin sözlü değerlendirmeye ilişkin performans ortalaması (Ort.= 3.53; SS=1.19), erkek öğrencilerin sözlü değerlendirmeye ilişkin performans ortalamasından (Ort.= 2.97; SS=1.25) daha yüksek çıkmıştır. Aynı şekilde sözsüz ölçüte ilişkin performans puanları arasında da cinsiyete göre anlamli bir farklılık bulunmuştur ( $t_{(253)} = 2,721; p < 0.5$ ). Kız öğrencilerin sözsüz değerlendirmeye ilişkin ortalaması (Ort.= 3.62; SS=1.25) yine erkek öğrencilerin sözsüz değerlendirmeye ilişkin ortalamasından (Ort.= 3.08; SS=1.28) daha yüksek çıkmıştır. Bu sonuçlar kızların hem sözlü hem de sözsüz performanslarının erkeklere kıyasla daha iyi olduğunu göstermektedir. Bir diğer ifadeyle ölçütlere ait performanslar üzerinde öğrencilerin cinsiyetlerinin etkili olduğu söylenebilir.

### **Öğrencilerin sözlü ve sözsüz ölçütlere ilişkin performansları arasında istatistiksel olarak anlamli bir ilişki var mıdır? Sorusun İlişkin Bulgular ve Yorum**

Öğrencilerin sözlü ve sözsüz ölçütlere ilişkin performansları iki uzman tarafından verilen puanların ortalaması alınarak hesaplanmıştır. İki uzmanın puan ortalamalarına göre sözlü ve sözsüz ölçütlere ilişkin performanslar arasındaki ilişki; her iki değişkene ait puanların normal dağılım göstermesi ve eşit aralıklı düzeyde olması nedeniyle Pearson Momentler Çarpımı Korelasyon Katsayısı ile belirlenmiştir. Değişkenler arasında pozitif yönde ( $r = .86; p < 0.05$ ) yüksek düzeyde bir ilişki olduğu görülmüştür. Bu bulgudan hareketle öğrencilerin sözlü performansları arttıkça sözsüz performanslarının da önemli oranda arttığı veya öğrencilerin sözlü performansları düştükçe sözsüz performanslarının önemli oranda düştüğü söylenebilir.

## TARTIŞMA ve SONUÇ

Rasch modeline göre yapılan analizlerden elde edilen bulgulara göre sözlü ve sözsüz ölçütlerin logit değeri -0,23 ve 0,23 ve logit değişim değeri 0,46 [0,23-(-0,23)] olarak; uyum içi ortalamaları  $M=0,99$  ve uyum dışı ortalamaları  $M=0,98$  olarak; ayırma indeksi ise 6,39 olarak elde edilmiştir. Bu bulgular ölçme aracının amaç doğrultusunda çalıştığını, geçerliğine ilişkin olumlu bulguları içerdiğini ve amaç doğrultusunda kullanılabileceğini öne sürmektedir. Orijinal kültüründe ölçme aracının kapsam geçerliği üzerine yapılmış çalışmalar ise ölçme aracının ölçülmek istenen özelliği ölçmeye yönelik geçerli bir enstrüman olduğunu kanıtlamıştır (Rutkowski, 1990, s. 92). Mevcut çalışmanın sonuçları ise ölçme aracının Türk kültüründeki geçerliğini ortaya koymaktadır.

Ölçme aracının güvenilirliğine ilişkin ise puanlayıcılar arası ve puanlayıcı içi güvenilirlik değerleri hesaplanmıştır. Buna göre puanlayıcılar arası güvenilirlik değerleri sözlü kriter için  $\alpha=.99$ , sözsüz kriter için  $\alpha=.98$  olarak elde edilmiştir. Sözlü ve sözsüz kriterler için puanlayıcı içi güvenilirlik değerleri ise birinci puanlayıcı için  $r=.94$  ve  $.98$ , ikinci puanlayıcı için ise  $r=.96$  ve  $.99$  olarak elde edilmiştir. “Şarkı Söyleme Tonunun Gelişimi Ölçeği” üçüncü versiyonunun orijinal güvenilirlik çalışmasında ise Pearson Momentler Çarpımı Korelasyon Katsayısı puanlayıcı arası ve puanlayıcı içi güvenilirlik değerlerinin belirlenmesi için kullanılmıştır. Buna göre ölçme aracının puanlayıcılar arası güvenilirlik katsayısının  $.80$  ve  $.85$  arasında, puanlayıcı içi güvenilirlik katsayısının ise  $.82$  ve  $.93$  arasında değiştiği tespit edilmiştir. Bu değerler ölçme aracının orijinal kültürü içindeki güvenilirliğini işaret etmektedir (Rutkowski, 1997, s. 206-207). Mevcut uyarlama çalışmasının sonuçları ölçme aracının Türk kültüründe de yüksek oranda güvenilirlik değerleri sağladığını göstermektedir.

Pereira ve Rodrigues (2019) Şarkı Söyleme Tonunun Gelişimi Ölçeğinin Portekiz kültüründeki geçerliği ve güvenilirliğini ele almıştır. Kapsam geçerliğine ilişkin alınmış uzman görüşlerinin sonuçları ölçme aracının Portekizli çocuklar üzerinde geçerli olduğunu göstermiştir. Ön- Şarkıcı (1), Tutarsız, konuşma ses bölgesinde şarkı söyleyen (1,5) ve konuşma ses bölgesinde şarkı söyleyen (2) ölçütlerinin görülme sıklığı görece düşük olsa da elde edilen sonuçların geneli ölçme aracının Portekiz kültüründeki geçerliliğini işaret etmektedir. Puanlayıcılar arası güvenilirlik katsayıları ise hem sözlü hem de sözsüz kategori için  $.95$  olarak elde edilmiştir. Bu çalışma ölçme aracının Portekiz kültüründeki geçerlik ve güvenilirliğini ortaya koymuştur. Chuang (2012) ölçme aracını Çin kültürüne uyarlamış ve orijinal ölçme aracındaki dokuz ölçütü genişleterek on bir ölçütlü bir yapı elde etmiştir. Mevcut çalışmanın sonuçları ise ölçme aracının dokuz ölçütlü yapısıyla Türk çocukları üzerinde geçerli ve güvenilir sonuçlar ürettiğini ortaya koymuştur.

Örneklemin şarkı söyleme tonu kullanım düzeyi, sınıf düzeyi değişkenine göre incelendiğinde 3. sınıfların sözsüz ölçütteki puanlarının, 2. sınıfların puanlarına kıyasla görece düşük olduğu belirlenmiştir; ancak bu fark istatistiksel olarak anlamlı değildir. Genel olarak sınıf düzeyi arttıkça şarkı söyleme tonu kullanım düzeyinin artma eğiliminde olduğu görülmektedir. Bu bulgu, literatürde biyolojik yaşın şarkı söyleme tonunun kullanımı üzerinde etkili olduğunu öne süren çalışmalar ile paralellik sağlamaktadır (Davies & Roberts, 1975; Goetze, 1985; Rutkowski & Miller, 2003a). Ancak biyolojik yaşın şarkı söyleme tonunun gelişimi ve şarkı söyleme becerileri üzerinde kayda değer bir etkisi olmadığını öne süren çalışmalar da bulunmaktadır (Mang, 2006; Levinowitz vd., 1998).

Literatürdeki bu çelişkili bulguların örneklem büyüklüğü, örneklem seçimi ve kültürel farklılıklardan ileri gelebileceği düşünülmektedir. Rutkowski & Miller (2003a)'ın ilkokul çocukları üzerinde yaptığı uzun dönem çalışmasının sonuçları, çocukların birinci sınıftan sonra şarkı söyleme becerileri üzerinde daha belirgin kazanımlar elde ettiğini ortaya koymuştur. Aynı çalışmanın sonuçları genel müzik eğitimi kapsamında uygulanan küçük grupla ve bireysel söyleme stratejileri ile beşinci sınıfın sonunda dek çoğu çocuğun en az başlangıç ses bölgesini kullanabilir düzeye geldiğini öne sürmektedir.

Cinsiyet değişkeni ele alındığında ise kız çocuklarının erkek çocuklara kıyasla Şarkı Söyleme Tonunun Gelişimi Ölçeği'nden daha yüksek puanlar aldığı belirlenmiştir. Bu farklılık kız çocuklarının erkek çocuklarına göre şarkı söyleme tonunu kullanmada avantaja sahip olduğunu öne sürmektedir. Literatür incelendiğinde kız çocuklarının lehine, cinsiyetin şarkı söyleme tonu kullanımı ve vokal doğruluk üzerinde etkili olduğunu işaret eden çalışmalar bulunmaktadır (Davies & Roberts, 1975; Goetze, 1985; Rutkowski, 1986; Guerrini, 2002, 2004; Mang, 2006). Mevcut çalışma bu çalışmaların bulguları ile paralellik göstermektedir. Rutkowski (1986, s. 243) doktora tezinde okulöncesi çocuklarının hem şarkı söyleme tonu kullanımında hem de gelişimsel müzik potansiyelinde kız çocuklarının lehine bir fark olduğunu tespit etmiştir; ancak bu fark istatistiksel olarak anlamlı değildir ve iki cinsiyet arasındaki fark oldukça azdır. Bu çalışmaların aksine cinsiyetin şarkı söyleme becerileri üzerinde etkili olduğuna ilişkin bulgulara ulaşamayan çalışmalar da mevcuttur (Smale, 1987; Welch vd., 1995; Welch vd., 1997; Cooper, 1995). Welch vd. (1997)'nin 5-7 yaş grubu üzerinde yaptığı üç yıllık uzun dönem çalışmasının sonuçları kız ve erkek çocuklarının vokal ses eşleştirme beceri düzeylerinin önceki çalışmaların aksine homojen olduğunu öne sürmektedir. Sadece 7 yaş grubunda şarkı seslendirme boyutunda kız çocuklarının lehine bir fark elde edilmiştir.

Ölçme aracının sözlü ve sözsüz ölçütleri arasında yüksek düzeyde anlamlı bir ilişki tespit edilmiştir. Bu öğrencilerin sözlü performansları arttıkça sözsüz performanslarının da önemli oranda arttığını veya öğrencilerin sözlü performansları düştükçe sözsüz performanslarının önemli oranda düştüğünü göstermektedir. Sözlü ve sözsüz ölçütlere ilişkin performans ortalamalarına bakıldığında ise öğrencilerin sözsüz ölçüte ilişkin performanslarının sözlü ölçüte ilişkin performanslardan görece daha yüksek olduğu tespit edilmiştir. Bu bulgu anasınıfı ve ilkokul üçüncü sınıf arasındaki çocukların şarkıları sözlerinden bağımsız olarak daha kolaylıkla ve netlikle seslendirdiğini öne sürmektedir. Literatürde şarkıların sözlü öğretimi veya sözsüz öğretiminin etkililiği üzerinde kesin bir uzlaşma olmasa da önemli ölçüdeki çalışmaların sonuçları sözlerden bağımsız olarak şarkı öğretimi gerçekleştiğinde çocukların şarkı söyleme tonunu daha iyi kullanabildiğini ve daha doğru şarkı söyleyebildiğini öne sürmektedir (Goetze, 1985; Welch vd., 1995; Rutkowski, 2013; Rutkowski & Miller, 2003b). Bu çalışmanın sonuçları ilgili literatür ile paralellik sağlamaktadır. Rutkowski (2013) şarkıların doğal hecelerle öğrenildiği enformel rehberlik temelli müzik öğrenme ortamında, okulöncesi çocukların şarkı söyleme tonu kullanımlarının kayda değer şekilde gelişim gösterdiğini tespit etmiştir. Welch vd. (1995) de beş yaş grubu üzerinde yaptıkları çalışmanın sonucunda, sözlerle maskelenmeyen müzikal öğelerde çocukların daha kesin vokal yanıtlar verdiğini gösteren bulgular elde etmiştir. Gelişmekte olan küçük şarkıcılar için sözlerin müzikal öğelerden önce gelebileceğini ve bu sebeple küçük çocuklara müzik ile sözlerin ayrı öğretilmesi gerektiğini öne sürmüşlerdir. Kendal (2018)'in elde ettiği sonuçlar ise

şarkı söyleme tonunun gelişimi ve ses doğruluğunun sağlanması arasında farklılık göstermektedir. Buna göre şarkıların öncelikle sözsüz olarak sunulmasının okul öncesi çocukların şarkı söyleme tonu gelişimini destekleyebileceğini, şarkıların metinle birlikte sunulmasının ise okul öncesi çocukların ses doğruluğunu destekleyebileceğini öne sürmektedir. Bu kanıtların aksine Levinowitz (1987) ve Smale (1987) çocukların şarkı söyleme kazanımları üzerinde sözlü müziklerin veya sözsüz müziklerin (natural hecelerin) kullanımının anlamlı bir fark oluşturduğuna ilişkin bulgular elde edememiştir.

Mevcut çalışmanın sonuçları Rutkowski tarafından uzun yıllar süren geliştirme ve revizyon çalışmaları sonucunda ortaya konulmuş olan “Şarkı Söyleme Tonunun Gelişimi Ölçeği”nin ana sınıfı-ilkokul 3. sınıf arasındaki Türk çocuklar üzerinde de geçerli ve güvenilir sonuçlar ürettiğini ortaya koymuştur. Türkçe geçerlik ve güvenilirliği kanıtlanmış bu ölçme aracının anasınıfı ve ilkokul döneminde genel müzik eğitimi derslerine giren sınıf öğretmenleri veya müzik öğretmenleri tarafından kullanılması, çocukların bireysel ihtiyaçlarının belirlenmesi ve bu doğrultuda öğretim programlarının tasarlanması hususunda önemli katkılar sağlayacaktır. Yanı sıra Türkçe geçerliği ve güvenilirliği kapsamlı olarak ele alınmış olan ölçme aracının, ulusal literatürdeki önemli bir boşluğu doldurarak ileriki araştırmalara kaynaklık edeceği düşünülmektedir.

## **ÖNERİLER**

Bu çalışmada Türkçe uyarlaması yapılmış, geçerliği ve güvenilirliği ortaya koyulmuş olan “Şarkı Söyleme Tonunun Gelişimi Ölçeği” nin, özellikle okul öncesi ve ilkokul sınıflarında kullanımına ve yaygınlaştırılmasına yönelik çalışmalar yapılması ve mevcut müzik eğitimi uygulamalarına ölçme aracı olarak dâhil edilmesi önerilmektedir. Bu çalışma, örneklemin anasınıfı ve ilkokul 3. sınıf arasında öğrenimine devam eden 253 çocuk olması ile sınırlıdır. Ölçme aracının geçerliği okul öncesi ve ilköğretimin tüm kademeleri dâhil edilerek daha geniş bir sınıf düzeyinde test edilebilir. Ayrıca Türk çocuklarının şarkı söyleme tonunun gelişimini farklı değişkenler açısından ölçmeye yönelik, farklı eğitim kademelerinde geniş örneklemlerle tarama çalışmalarının yapılması da önerilmektedir. Durum tespitine yönelik bu tür çalışmaların, müzik eğitimi sistemimizdeki mevcut eksikliklerin giderilmesi ve iyileştirilmesi hususunda yol gösterici olacağı düşünülmektedir.

## **Etik Metni**

Bu makalede dergi yazım kurallarına, yayın ilkelerine, araştırma ve yayın etiğine, dergi etik kurallarına uyulmuştur. Makaleyle ilgili olarak ortaya çıkabilecek ihlallerin sorumluluğu yazar(lar)a aittir. Makale için etik kurul izni Gazi Üniversitesi Etik Komisyonu'nun 10.03.2023 tarih ve 2023- 248 sayılı kararı ile alınmıştır.

**Yazar(lar)ın Katkı Oranı Beyanı:** Her iki yazar da makaleye eşit katkıda bulunmuştur. Yazarlar makalenin İngilizce dil denetimi için Gazi Üniversitesi Akademik Yazma Uygulama ve Araştırma Merkezine teşekkürlerini sunmaktadır.

**KAYNAKÇA**

- Aaron, J. C. (1990). *The effects of vocal coordination instruction on the pitch accuracy, range, pitch discrimination, and tonal memory of inaccurate singers* [Unpublished doctoral dissertation]. University of Iowa.
- Abril, C. R. (2016). Untangling general music education: concept, aims, and practice. In C. R. Abril & B. M. Gault (Eds.). *Teaching general music: Approaches, issues, and viewpoints*. (pp. 5-22). Oxford University Press.
- Bilen, S. (1995). *İşbirlikli Öğrenmenin Müzik Öğretimi ve Güdüsel Süreçler Üzerindeki Etkileri* [Yayımlanmamış Doktora tezi]. Dokuz Eylül Üniversitesi.
- Boardman, E. L. (1964). *An investigation of the effect of preschool training on the development of vocal accuracy in young children* [Unpublished doctoral dissertation]. University of Illinois.
- Chuang, M. J. (2012). Construction of a children's singing voice measure. *Problems in Music Pedagogy*, 11, 93-101.
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale:Lawrence Erlbaum Associates.
- Cooper, N. A. (1995). Children's singing accuracy as a function of grade level, gender, and individual versus unison singing. *Journal of Research in Music Education*, 43(3), 222-231. <https://doi.org/10.2307/3345637>
- Davies, A. D. M., & Roberts, E. (1975). Poor Pitch singing: a survey of its incidence in school children. *Psychology of Music*, 3(2), 24-36. <https://doi.org/10.1177/030573567532004>
- Değer, A. Ç. (2012). *Çocuk korolarının eğitiminde bir yaklaşım olarak eğitsel oyun kullanımının öğrencilerin müziksel erişimi düzeylerine etkisi* [Yayımlanmamış Doktora tezi]. Gazi Üniversitesi.
- Dittemore, E. E. (1968). *An Investigation Of Some Musical Capabilities Of Elementary School Students* [Unpublished doctoral dissertation]. University of Iowa.
- Doneski, S. M. (2005). *The effects of wait time on the tonal pattern performance accuracy of second and fourth-grade students*. [Unpublished doctoral dissertation]. University of Hartford.
- Goetze, M. (1985). *Factors affecting accuracy in children's singing (music, Voice)*. [Unpublished doctoral dissertation]. University of Colorado.
- Gordon, E. E. (2007). *Awakening new-borns, children, and adults to the world of audiation*. GIA.
- Gordon, E. E. (2013). *Music Learning Theory for new-born and young children*. GIA.
- Gould, A. O. (1968). *Finding and learning to use the singing voice: a manual for teachers*. Eric Pub.
- Gould, A. O. (1969). Developing Specialized Programs for Singing in the Elementary School. *Bulletin of the Council for Research in Music Education*, 17, 9-22.
- Guerrini, S. C. (2002). *The acquisition and assessment of the developing singing voice among elementary students* [Unpublished doctoral dissertation]. Temple University
- Guerrini, S. C. (2004). The relationship of vocal accuracy, gender, and music aptitude among elementary students. *Visions of Research in Music Education*, 4(1), 4.
- Guerrini, S. C. (2006). The developing singer: Comparing the singing accuracy of elementary students on three selected vocal tasks. *Bulletin of the Council for Research in Music Education*, 21-31.

- Gürgen, E. T. (2007). *Orff-Schulwerk ve Kodaly yöntemi'nin vokal doğaçlama, müziksel işitme ve şarkı söyleme becerileri üzerindeki etkileri* [Yayımlanmamış Doktora tezi]. Dokuz Eylül Üniversitesi.
- Hale, M. R. (1976). *An experimental study of the comparative effectiveness of harmonic and melodic accompaniment in singing as it relates to the development of a sense of tonality* [Unpublished doctoral dissertation]. The State University of New York.
- Hall, C. (2005). Gender and boys' singing in early childhood. *British Journal of Music Education*, 22(1), 5-20. <https://doi.org/10.1017/S0265051704005960>
- Hayes, A. F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1:1, 77-89, DOI: [10.1080/19312450709336664](https://doi.org/10.1080/19312450709336664)
- Hornbach, C. M., & Taggart, C. C. (2005). The relationship between developmental tonal aptitude and singing achievement among kindergarten, first-, second-, and third-grade students. *Journal of Research in Music Education*, 53(4), 322. <https://doi.org/10.2307/3648430>
- Jaffurs, S. E. (2000). *The relationship between singing achievement and tonal music aptitude* [Unpublished master's thesis]. Michigan State University
- Karaönçel, F. (2021). *Sanal ve akustik çalgı eşlikli şarkı öğretimi uygulamalarının müziksel işitme ve şarkı söyleme becerilerine etkisi* [Yayımlanmamış Doktora tezi]. İnönü Üniversitesi.
- Kendal, J. L. (2018). *The effects of with-text and without-text song presentation styles on preschoolers' singing voice use and pitch accuracy* [Unpublished doctoral dissertation]. University of Maryland.
- Kılıç, S. (2014). Etki büyüklüğü. *Journal of Mood Disorders*, 4 (1), 44-6. DOI: 10.5455/jmood.20140228012836
- Kim, S. (2000). *The effects of sequential movement activities on first-grade students' solo singing abilities* [Unpublished master's thesis]. The University of Southern Mississippi
- Leighton, G., & Lamont, A. (2006). Exploring children's singing development: Do experiences in early schooling help or hinder? *Music Education Research*, 8(3), 311-330. <https://doi.org/10.1080/14613800600957461>
- Levinowitz, L. M. (1987). *An experimental study of the comparative effects of singing songs with words and without words on children in kindergarten and first grade* [Unpublished doctoral dissertation]. Temple University.
- Lewinowitz, L. M. Barnes, P. Guerrini, S. Clement, M. D'April, P. Morey, M. J. (1998). Measuring singing voice development in the elementary general music classroom. *Journal of Research In Music Education*, 46 (1), 35-47.
- Mang, E. (2006). The effects of age, gender, and language on children's singing competency. *British Journal of Music Education*, 23(2), 161-174. <https://doi.org/10.1017/S0265051706006905>
- Mizener, C. P. (2008). Our singing children: developing singing accuracy. *General Music Today*, 21(3), 18-24. <https://doi.org/10.1177/1048371308317086>
- Nakamura, Y. (2002). Teacher assessment and peer assessment in practice. *Educational Studies*, 44, 203-215.
- Newlin, G. A. (2004). *The effects of part-work instruction on first-grade part-singing acquisition and achievement*. Unpublished doctoral dissertation]. University of Hartford.



- Nichols, B. E. (2017). Constructing singing assessments for the music classroom. *General Music Today*, 30(3), 13-17. <https://doi.org/10.1177/1048371317690864>
- Pereira, A. I., & Rodrigues, H. (2019). The relationship between Portuguese children's use of singing voice and singing accuracy when singing with text and a neutral syllable. *Music Perception*, 36(5), 468-479. <https://doi.org/10.1525/mp.2019.36.5.468>
- Rutkowski, J. (1986). *The effect of restricted song range on kindergarten children's use of singing voice and developmental music aptitude (elementary, preschool)*. [Unpublished doctoral dissertation]. State University of New York
- Rutkowski, J. (1997). The nature of children's singing voices: Characteristics and assessment. *The phenomenon of singing*, 1, 201-209.
- Rutkowski, J., & Snell Miller, M. (2003a). A longitudinal study of elementary children's acquisition of their singing voices. *Update: Applications of Research in Music Education*, 22(1), 5-14. <https://doi.org/10.1177/87551233020220010901>
- Rutkowski, J., & Miller, M. S. (2003b). The effectiveness of frequency of instruction and individual/small-group singing activities on first graders' use of singing voice and developmental music aptitude. *Contributions to Music Education*, 30 (1), 23-38.
- Rutkowski, J., & Miller, M. S. (2003c). The effect of teacher feedback and modeling on first graders' use of singing voice and developmental music aptitude. *Bulletin of the Council for Research in Music Education*, 1-10.
- Rutkowski, J. (2013). Kindergarten children's use of singing voice achievement and developmental music aptitude in the context of informal music instruction. *Perspectives: Journal of the Early Childhood Music & Movement Association*, 8(1), 6-11. [https://doi.org/10.1386/IJMEC\\_0225\\_1](https://doi.org/10.1386/IJMEC_0225_1)
- Rutkowski, J. (2015). The relationship between children's use of singing voice and singing accuracy. *Music Perception*, 32(3), 283-292. <https://doi.org/10.1525/mp.2015.32.3.283>
- Rutkowski, J. (1990). The measurement and evaluation of children's singing voice development. *The Quarterly*, 1(1-2), 81-95.
- Salvador, K. (2010). How can elementary teachers measure singing voice achievement? a critical review of assessments, 1994-2009. *Update: Applications of Research in Music Education*, 29(1), 40-47. <https://doi.org/10.1177/8755123310378454>
- Smale, M. J. (1987). *An investigation of pitch accuracy of four- and five-year-old singers*. [Unpublished doctoral dissertation]. University of Minnesota
- Smith, J. (2006). Every child a singer: techniques for assisting developing singers. *Music Educators Journal*, 93(2), 28. <https://doi.org/10.2307/3878468>
- Steenwyk, L. A. (2004). *An investigation of the use of singing games in measuring singing voice development in young children*. [Unpublished master's thesis]. Michigan State University.
- Tavani, F. (1998). *The relationships among first grade students' tonal aptitude, singing vocal development, and developmental age*. [Unpublished master's thesis]. Rowan University.

- Tu, C. M. (2009). Music aptitude, tonal accuracy achievement, and singing voice development of kindergarten and first-grade children. *Research Perspectives in Music Education*, 13(1), 29-37.
- Uçal, E. (2003). Okulöncesi müzik eğitiminde Orff öğretisinin müziksel beceriler üzerindeki etkileri [Yayımlanmamış Yüksek Lisans tezi]. Dokuz Eylül Üniversitesi.
- Uysal, G. (2004). *İlköğretimde işbirlikli öğrenmenin müzik öğretiminde sınıf atmosferi ve şarkı söyleme becerileri üzerindeki etkisi* [Yayımlanmamış Yüksek Lisans tezi]. Dokuz Eylül Üniversitesi.
- Vande Wege, R. M. (2005). *The effect of tonal pattern instruction on the singing voice development of first grade students*. [Unpublished master's thesis]. Michigan State University.
- Welch, G. F. (1979). Poor pitch singing: A review of the literature. *Psychology of Music*, 7(1), 50-58. <https://doi.org/10.1177/030573567971006>
- Welch, G. F. (1985). A schema theory of how children learn to sing in tune. *Psychology of Music*, 13(1), 3-18. <https://doi.org/10.1177/0305735685131001>
- Welch, G. F. (1986). A developmental view of children's singing. *British Journal of Music Education*, 3(3), 295-303. <https://doi.org/10.1017/S0265051700000802>
- Welch, G. F., Sergeant, D. C., & White, P. J. (1995). The singing competencies of five-year-old developing singers. *Bulletin of the Council for Research in Music Education*, 155-162.
- Welch, G. F., Sergeant, D. C., & White, P. J. (1997). Age, sex, and vocal task as factors in singing "in tune" during the first years of schooling. *Bulletin of the Council for Research in Music Education*, 153-160.
- Welch, G. F. (2002). Early childhood musical development. L. Bresler & C. M. Thompson (Ed.), *The arts in children's lives* (ss. 113-128). Springer Netherlands. [https://doi.org/10.1007/0-306-47511-1\\_10](https://doi.org/10.1007/0-306-47511-1_10)
- Welch, G. F. (2016). Singing and vocal development. G. McPherson (Ed.), *The child as musician*, (pp.311-329). Oxford University Press.
- Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance, and self-assessment. *Musicae Scientiae*, 12(1), 3-26. <https://doi.org/10.1177/102986490801200102>