VALIDITY AND RELIABILITY STUDY OF THE LISTENING COMPREHENSION SCALE

© Pınar GIRMEN (Eskisehir Osmangazi University, Eskisehir, Turkey)

pgirmen@ogu.edu.tr

© Emel BAYRAK (Milli Egitim Bakanlıgı, Eskisehir, Turkey)

emelbayrak84@hotmail.com

Language is both the source and reflection of human thinking. Listening, which is the first experience with the mother tongue in a child's life, acts as a means for learning other language skills as well as being the most used language skill in an individual's social and academic life. An effective learning experience requires that students move from where they already are. This change can only be determined by standardized scales structured in harmony with the nature of the variable to be measured. Those scales for determining listening level can involve students in the active nature of listening by means testing situations they present in practical process. The primary aim of this study is to present the validity and reliability test results of the listening scale developed in order to determine the primary school 5th graders' listening comprehension skill level. The scale's validity and reliability analyses were applied in Eskişehir city centre with a total of 700 primary school 5th graders chosen through stratified sampling. The scope of the listening scale comprised the listening comprehension acquisitions covered in the Primary Education Turkish course curriculum. Aimed at measuring mental skills such as understanding, predicting, summarizing, classifying, ordering, concluding, distinguishing and comparing what is listened to, the scale was designed as fifteen separate parts. Each part consisted of instructions, passages and questions. The draft version of the scale was applied twice by the researcher and once by two expert academics being revised each time and semi-structured interviews were conducted with the students. For times when the students might be distracted, enjoyable listening activities called "leisure time" were designed in order to eliminate the validity problems to be caused by the implementation process. Relevant field experts were consulted regarding the scope and face validity of the scale. The item score matrix, item discrimination power and item difficulty index of the scale items were calculated as well as other statistics such as the scale's arithmetic mean, standard deviation, validity and reliability. The item total score correlation, explaining the scale's internal validity and item discrimination index, was calculated with the Pearson Product-moment Correlation Coefficient and Item Discrimination Power Based on Group Differences. The CronbachAlpha Coefficient was used in determining the reliability of the scale. The findings from the analyses revealed that the scale is a valid and reliable instrument.

Keywords: Listening, listening scale, primary education

Language is both the source and reflection of human thinking. This feature of language constitutes the basic dynamics on which the individual's social and academic life is to be built. In the deliberate domestication process which they desire, educational institutions are expected to improve individuals' skills of recognizing, developing and using these functions of language. Language has the power to create, convey, influence and manipulate intellectual and spiritual worlds of individuals and societies.

Since language is a system which involves mind, culture and universe, mother tongue education is built upon interactive basic language skills each of which requires distinct acquisitions. These skills are divided into two as "comprehension" and "expression". Comprehension involves "reading and listening" while expression involves "speaking and writing" skills of language.

Contemporary advancements in communication technology require that people continuously access new information. It is only through receiving a proper listening education that people can approach new information with a sense of awareness (Doğan, 2007:1). Nonetheless, speaking and listening skills are covered less than reading and writing in formal education (Ergin & Birol, 2000:114). There is a limited number of studies conducted into listening comprehension in comparison with those studies in other skills (Girmen, Kaya & Bayrak, 2010:135). However, children learn their mother tongues by hearing it and listening skill can become a more efficient way of comprehension by means of systematic approaches that are appropriate for its nature (Kalaycı & Temur, 2005:55).

As the only means of comprehension until the school age for people, who start language learning process with listening in prenatal period, listening improves and forms the base for other language skills as well as knowledge, feeling and thought universe of people during that period (Sever, 2010:9) (Özbay, 2005:9; Güneş, 2007:73). Because most relationships of people are based on telling and listening, listening skill is a major means of obtaining information, learning and comprehension (Özbay, 2005).

Physiologically, while hearing is a process in which sound waves turn into mechanical vibrations in the middle ear and nerve currents reaching the brain in the inner ear in their travel from the outer ear to the mind (Ergin & Birol, 2000:115), listening is an activity in which the individual gets involved in the process cognitively, affectively and kinetically. In listening environments, individuals are not passive recipients, instead they are receivers that use their mental skills to create meaning (Ungan, 2007:149; Karadüz, 2010:42).

Mental processes are important in the formation of meaning and concentration in listening and claims, that listening is an activity that involves much more than hearing. Similarly, Johnson (1951:13) defines listening as the ability to understand and respond actively.

Listening is not just allowing sound waves to enter into ears and it requires more effort like the way reading is more than merely looking at written words (Temur, 2010:306). Therefore, listening gets beyond hearing sound waves and involves perceiving the meanings created by those waves,

reflecting upon these meanings and taking action when necessary (Doğan, 2008:263).

Research indicates that most people get involved in situations of listening more than speaking (Ergin & Birol, 2000:114). The traditional dependence of teaching on lecturing causes the time allocated for listening in classrooms to reach 83% (Özbay, 2005:13). Students spend 50% of their school time in primary school and about 90% of that in university for listening activity (Özbay, 2005:124). Person devotes 42% of their time when they are with other people to listening, students listen to their teachers and friends for 2.5-4 hours a day at school and academic achievement is strongly associated with students' listening skill (in Sever, 2010:9). The table below shows findings from research conducted into daily periods of time devoted for various language skills (Temur, 2010:306).

Author	Sample	Reading	Writing	Speaking	Listening
Werner,1975	Student- workers	.13	.08	.23	,55
Berker et al.,1980	University Students	.17	.14	.16	.53
Perras&Weitzel,1981	University Students	.14	.08	.03	.15
Bohlekn,1999	University Students	.13	.12	.22	.53
Davis,2001	University Students	.12	.10	.31	.34
Janusik&Wolvin,2007	University Students	.07	.08	.20	.23

Table 1. Daily periods of time devoted for various language skills

The spontaneous acquisition of listening skills and habits, which surround the relations in every facet of life, in the pre-school period does not necessarily mean that an individual's listening skill possesses higher-order cognitive skills. Having students acquire these skills and habits is a responsibility of mother tongue education (Sever, 2010:9). When listening skill is thought to be acquired automatically and it is not taught systematically, a problem of quality can occur and this, in turn, can bring about many social problems as well as communication problems.

This situation indicates the necessity to design learning and teaching situations that are appropriate for having students acquire behavioural patterns about the cognitive, affective and kinetic aspects of listening skill within mother tongue teaching activity (Sever, 2010:9). Since human beings learn their mother languages mainly by listening to people around them, the primary way of studying in Turkish language course at school is listening in accordance with this natural way of learning. Listening is an important component performed in a natural course in classroom activities (Cemiloğlu, 2004:97).

Teaching listening involves activities designed to make individuals be aware about understanding and making sense of what they hear and to have them gain the ability to use their knowledge and energy properly. Listening education is aimed at raising students who know what to listen to and why to listen to something; can concentrate on what is told, spoken or read; explore comprehension situations; and make attempts to understand and analyze them (Koç & Müftüoğlu, 1998:55).

As mentioned above, despite its significance, listening education is an area which has not been given the necessary importance in terms of both

application system and assessment tools. The limited learning experiences caused by these inadequacies cause students fail to use listening as an active skill in their academic and social lives. Listening education is needed for raising individuals who are equipped with listening skills and have the habit of thinking, inquiring and communicating efficiently. In the same way, studies to be conducted into listening skill education are needed so that listening education can reach the desired level. In order for the impact created by the individuals who are able to listen to expand into families, schools and finally whole society, the quality of the education given in educational institutions should be assessed by means of standardized tools of measurement and the whole process should be restructured in light of the findings to be obtained.

The aim of this study is to present the validity and reliability test results of the listening comprehension scale developed as an instrument in order to measure the primary school 5th graders' listening comprehension skill level.

Method

The study population consisted of primary school 5th graders in the primary schools in Eskişehir city centre. Research data were collected from a total of 700 5th graders attending 14 randomly-chosen primary schools in Eskişehir city centre in 2010–2011 school year. The schools and the number of the participants are shown in Table 2.

No	Name of the School	Number of Participants	Percentage
1	Av. Mail BÜYÜKERMEN P.S.	27	0,0385
2	Dr. Halil AKKURT P.S.	62	0,0885
3	Ticaret Odası P.S.	58	0,0828
4	Cemalettin SARAR P.S.	45	0,0642
5	Sami SİPAHİ P.S.	44	0.0628
6	Şehit Ali Gaffar OKKAN P.S.	80	0,1142
7	Kardeşler P.S.	50	0,0714
8	İbrahim KARAOĞLANOĞLU P.S.	66	0,0942
9	Orgeneral Halil SÖZER P.S.	36	0,0514
10	Dumlupınar P.S.	40	0,0571
11	Sinan ALAĞAÇ P.S.	51	0,0728
12	Şehit Subutay ALKAN P.S.	24	0,0342
13	Seker P.S.	77	0,1100
14	j İki Eylül P.S.	40	0,0571
	TOTAL	700	

Table 2. Distribution of students according to schools in the study

Problem Definition

The listening comprehension scale was aimed at determining the primary school 5th graders listening comprehension skill levels. Therefore, the listening comprehension scale comprised the listening comprehension acquisitions included in the primary education Turkish language course curriculum.

It was determined that among 35 acquisitions, 9 acquisitions required long-term observations to be measured. Therefore, only the structures to be used in measuring the remaining 26 acquisitions were designed.

Item Writing

In the scale development process, the isomorphic questions which are appropriate for the nature of the acquisitions the scale was intended to measure and the passages to place these questions were created. During the creation of texts, this age group's attention span and interests were taken into consideration in determining the ideal passage length and content. The scale aimed at measuring mental skills such as comprehending, predicting, summarizing, categorizing, ordering, making decisions about what is listened to.

The scale was designed in fifteen separate parts designed according to the features of the skill domain it was intended to measure. Each part of the scale started with instructions that informed the students about the corresponding listening objective. Each of the parts consisted of instructions, the questions aimed at measuring an acquisition and a listening passage.

Determining the time allowed for the scale. The time allowed to students for implementation of the test is particularly significant because, unlike reading, it is not possible for students to go back whenever they want while applying the listening scale. Therefore, the scale was individually applied with some 5th graders with different academic levels in order to calculate the ideal application time of the scale; the time spent by the students to answer the parts was measured for each student and then the average application times were determined for each part.

Scale instructions. Instructions were prepared and placed at the top of the scale to make sure that the students would understand and answer the scale easily. The instructions presented information about how many items there were in the scale, the aim and duration of the scale and what to do with the scale.

Receiving Expert Opinions and Creating Trial Implementation Form

First of all, the answer to the question "To what extent are the items in the scale efficient in covering and collecting factual and/or speculative data desired?" is sought. Experts are referred to obtain an answer to this question about the content validity of the scale (Büyüköztürk, 2008:134). Therefore, with the participation of the researcher, expert opinions were received in order to determine the content and face validity of the scale. Also, experts working in primary school Turkish language education area examined suitability of the application duration for that age group, relevance of the scale items to the corresponding acquisitions and the items' competency in determining the students' listening comprehension scales.

In order to obtain field teachers' opinions about the scale, a few field teachers were chosen and given the scale and "the scale assessment form" to evaluate it. The teachers were also given the opportunity to express their opinions apart from the form. The scale was revised in light of the feedbacks from the field experts and teachers.

Trial Implementation

In order to determine the applicability of the scale in a classroom environment, the draft version of the scale, which consisted of instructions, passages and questions, was initially applied in a 5th grade classroom having students with medium level socio-economic background. The students were monitored during this implementation and after that, their opinions about the scale were determined by means of semi-structured interviews. The scale was then revised in light of the feedbacks and applied with students from another class.

After that, the scale was applied by two field expert academics, in a school of low socio-cultural level and was revised once more. Following this last trial implementation, it was decided that the scale would be divided into two 20-minute parts and applied in two separate class hours. During the rest of a 40-minute class hour, on the other hand, it was decided that students would spend their time with enjoyable listening activities called "leisure time".

All these actions were taken in order to ensure the validity of the test. After the implementation – the durations in terms of the distribution of the parts within the scale was determined.

Validity and Reliability Tests

The statistics describing the distribution of the responses to the items were used in validity and reliability analyses of the scale. Opinions of field experts were referred as a part of the content and face validity tests of the scale. Item statistics such as the discrimination power of each item, mean, frequency, percentage and standard deviation were calculated. Arithmetic mean, standard deviation and variance calculations were carried out for the overall total scores of the scale items and its view according to normal distribution curve was summarized graphically. During the calculation of the statistics about difficulty of the scale, the validity and reliability of the scale, "Cronbach Alpha Coefficient" was used to determine the internal consistency of the scale and "Pearson Product-moment Correlation Coefficient-Item Total Score Correlation- and Item Discrimination Power Based on Group Differences" was calculated to determine discrimination power of the items.

Findings Findings about the Validity and Reliability Tests of the Listening Scale

Internal Consistency Tests (Cronbach Alpha Coefficient). The Cronbach Alpha reliability coefficient was found as alpha = .7311 and the standardized item was found as alpha = .8801 for the scale reduced to 29 items as a result of the validity test carried out with a total of 700 students. This reliability coefficient is regarded to possess a high level of reliability in educational and social sciences.

Item Total Correlation (Pearson Product-Moment Correlation Coefficient). A higher score in item total correlation, which explains the relationship between the scores received from the test items and the total score of the test, indicates that the items represent similar behaviours and the test has a high level of internal consistency; in general, it could be suggested

that items with .30 or higher item total correlation values discriminate well among subjects (Büyüköztürk, 2011:171).

Table 3. Item Total Correlation of the Scale Items (Pearson Correlation Coefficient)

Question Number	1	2	3	4	5	6	7	8	9	10
Pearson Correlation	,322	,509	,598	,547	448	,428	,475	,359	,578	,553
Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
N	700	700	700	700	700	700	700	700	700	700
Question Number	11	12	13	14	15	16	17	18	19	20
Pearson Correlation	,417	,411	,505	,594	,472	,343	,451	,497	,448	,389
Sig. (2- tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
N	700	700	700	700	700	700	700	700	700	700
Question Number	21	22	23	24	25	26	27	28	29	g-t
Pearson Correlation	,506	,436	,398	,422	,335	,351	,479	,351	,361	1
Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
N	700	700	700	700	700	700	700	700	700	700

Correlation is significant at the 0,01 level

As can be seen in the table, the item total correlation varied between .322 and .598, which is above the .30 value required for the item total correlation. Therefore, it could be suggested that the scale items highly discriminate among the participants.

Item Discrimination Power Based on Group Differences (Upper – Lower Group). This analysis was carried out to determine whether the scale items distinguished between the subjects with and without the quality measured. The correlation values varied between 386 and .687, which can be interpreted to show that the scale items had the power to discriminate between those knowing and those not knowing.

Table 4. Item Discrimination Power Based on Group Differences among the Scale Items

Question Number	Top 27- Down 27 Correlation	Mean Top 27	Mean Down 27	Std. Deviation top 27	Std. Deviation Down 27
1	,444	,7672	,3228	,42374	,46877
2	,661	,8571	,1958	,35086	,39784
5	,539	,8783	,3386	,32780	,47450
6	,534	,8307	,2963	,37602	,45784
7	,582	,8148	,2328	,38948	,42374
8	,481	,6931	,2116	,46242	,40956
10	,687	,9153	,2275	,27911	,42034
11	,539	,7619	,2222	,42705	,41684
12	,569	,6772	,1534	,46877	,36137
13	,650	,8942	,2434	,30842	,43027
15	,579	,7884	,2116	,40956	,40956
16	,386	,8571	,4709	,35086	,50048
17	,507	,9471	,4392	,22445	,49760
18	,571	,8730	,3016	,33384	,46017
19	,508	,9101	,4021	,28687	,49163
20	,455	,8783	,4233	,32780	,49539
21	,640	,8624	,2222	,34536	,41684
22	,492	,9471	,4550	,22445	,49930
23	,433	,8889	,4550	,31510	,49930
27	,555	,9153	,3598	,27911	,48121
29	,439	,6667	,2275	,47266	,42034

Content Validity Tests. Content validity deals with whether the items in a scale are appropriate for the purpose and represent the domain which is intended to be measured. To this end, experts were asked examine and give opinions about the scale. A field expert and the researchers examined each item in terms of the acquisition it is intended to measure and revised those not serving their objectives.

Determination of Item Statistics. The frequency gives information about the difficulty of the item by showing the number persons responding to the item. The frequency, percentage, mean and standard deviations of the scale items are shown below.

Table 5. Item Statistics of the Scale Items

Question Number	Valid	f	%	p	r
Q.1	,00	323	46,1	.54	.50
	1,00	377	53,9		
Q.2	,00	335	47,9	.52	.50
	1,00	365	52,1		
Q.3	,00	78	11,1	.74	.36
	,25	45	6,4		
	,50	137	19,6		
	,75	1	,1		
	1,00	439	62,7		
Q.4	,00	15	2,1	.80	.26
	,25	46	6,6		
	,50	89	12,7		
	,75	172	24,6		
	1,00	378	54,0		
Q.5	,00	257	36,7	.63	.48
	1,00	443	63,3		
Q.6	,00	295	42,1	.58	.49
	1,00	405	57,9		
Q.7	,00	324	46,3	.54	.50
	1,00	376	53,7		
Q.8	,00	403	57,6	.43	.50
	1,00	297	42,4		
Q.9	,00	57	8,1	.70	.36
	,25	93	13,3		
	,50	166	23,7		
	1,00	384	54,9		
Q.10	,00	285	40,7	.60	.50
	1,00	415	59,3		
Q.11	,00	371	53,0	.47	.50
	1,00	329	47,0		
Q.12	,00	434	62,0	.38	.49
	1,00	266	38,0		
Q.14	,00	34	4,9	.65	.31
	,25	121	17,3		
	,50	153	21,9		
	,75	185	26,4		
	1,00	207	29,6		
Q.15	,00	352	50,3	.50	.50
	1,00	348	49,7		
Q.16	,00	225	32,1	.68	.47
	1,00	475	67,9	-	-
Q.17	,00	208	29,7	.70	.48
Ç	1,00	492	70,3		
Q.18	,00	292	41,7	.59	.50
Ç,	1,00	408	58,3		
Q.19	,00	209	. , -		
	1,00	491			

Q.20	,00	230	32,9	.68	.47
	1,00	470	67,1		
Q.21	,00	327	46,7	.53	.50
	1,00	373	53,3		
Q.22	,00	207	29,6	.70	.46
	1,00	493	70,4		
Q.23	,00	230	32,9	.67	.47
	1,00	470	67,1	.45	.35
Q.24	,00	141	20,1		
	,25	179	25,6		
	,50	223	31,9		
	,75	1	,1		
	1,00	156	22,3	.75	.27
Q.25	,00	9	1,3		
	,25	39	5,6		
	,50	133	19,0		
	,75	289	41,3		
	1,00	229	32,7		
Q.26	,00	8	1,1	.81	.23
	,25	32	4,6		
	,50	91	13,0		
	,75	213	30,4		
	1,00	356	50,9		
Q.27	,00	230	32,9	.67	.47
	1,00	470	67,1		
Q.28	,00	57	8.1	.72	.33
	,25	62	8.9		
	,50	135	19,3		
	,75	113	16,1		
	1,00	333	47,6		
Q.29	,00	379	54,1	.46	.50
	1,00	321	45,9		

The analysis performed with SPSS 11.5 revealed that among the frequency and percentage figures relating to the correct or incorrect answers to the scale questions given by a total of 700 5th grade students with whom the listening comprehension scale was applied, the frequency of correct answers varied between 490 and 266 while the frequency of incorrect answers varied between 434 and 210. These results indicate that the scale was of medium difficulty, which was a desired quality.

The highest possible value that a question can have is "1". The mean of any question can receive values in 0-1 interval. It was determined that the mean of the questions varied between .3809 and .8135 and their standard deviation varied between .23199 and .50522. The frequency distributions of the scale questions in normal distribution curve are shown in the graphic below.

S_GT

100

60

40

Std. Dev = 5.68
Mean = 17.8
N = 701.00

Graphic 1. Distribution of the Item Overall Totals in Normal Curve

Item Overall Total Statistics. It was determined that the lowest score in the sample was 2,75 while it was 29 points. This shows that the scale scores displayed a wide distribution. The fact that the mean of the scale was 17,7507 could be regarded as an indication of the fact that the scale had a medium level of difficulty.

S GT

N Variance Minimum Std. Maximum Mean Deviation 700 29 2,75 17,7507 32,106 S gt 5,66625 700 Valid N

Table 6. Item Overall Total Statistics

Discussion and conclusion

Can the fact that listening starts as early as the prenatal period and it is the most used language skill ensure that individuals are able to make sense of the intensive verbal content of their social and academic lives? It is essential that listening skill should be approached with a systematic understanding of education which is appropriate for its nature. In constructivism, on which Turkish primary education program over the last years has been built, the source of learning experiences is students' previous learning experiences. Considering the fact that listening is a skill learnt in the pre-school period, educational institutions should identify students' incomplete or improper learning and provide them with learning experiences which are relevant to their areas of interest. Standardized tools of measurement are required to carry out these determinations. The "Listening Comprehension Scale" developed in this study is aimed at measuring students' skills at various cognitive levels in terms of listening comprehension. The observations during the implementation of the scale showed that the students realized the nature of listening and concentrated more on the tasks due to the test situation following listening. In this regard, the measurement tools of listening comprehension skill can also be used as a useful learning tool during teaching.

The findings from validity and reliability analyses conducted on the listening comprehension scale prepared in accordance with the Primary

Education 5th Grade Turkish Language Program acquisitions revealed that this scale possessed applicable qualities.

References

BÜYÜKÖZTÜRK, Ş. (2008). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi Yayınları.

BÜYÜKÖZTÜRK, Ş. (2011). *Veri analizi el kitabı*. Ankara: Pegem Akademi Yayınları. CEMILOĞLU, M. (2004). *Türkçe öğretimi*. İstanbul: Aktüel Yayıncılık.

DOĞAN, Y. (2007). İlköğretim ikinci kademede dil becerisi olarak dinlemeyi geliştirme çalışmaları. Yayınlanmamış doktora tezi, Gazi Universitesi, Ankara.

Doğan, Y. (2008). İlköğretim yedinci sınıf öğrencilerinin dinleme becerilerini geliştirmede etkinlik temelli çalışmaların etkililiği. *Turkish Educational Sciences Journal*, 6(2), 261-286.

ERGIN, A., & BIROL, C. (2000). Eğitimde iletişim. Ankara: Anı Yayıncılık.

GIRMEN, P., KAYA, F., & BAYRAK E. (2010). Türkçe eğitiminde yaşanan sorunların lisans üstü tezlere dayalı olarak belirlenmesi. 9. Ulusal Sınıf Öğretmenliği Sempozyumu, 133-138.

GÜNEŞ, F. (2007). *Türkçe öğretimi ve zihinsel yapılandırma*. Ankara: Nobel Yayınları.

JOHNSON, O. K. (1951). The effect of classroom training upon listening comprehension. *The Journal of Communication*. I, 58.

KALAYCI, N., & TEMUR, T. (2005). İlköğretim okullarında dinleme becerisini nasıl geliştirebiliriz. *Eğitime Bakış*, 53-58.

KARADÜZ, A. (2010). Türkçe ve sınıf öğretmeni adaylarının dinleme stratejilerinin değerlendirilmesi. *Sosyal Bilimler Enstitüsü Dergisi*, 29, 40-55.

KOÇ, P., & MÜFTÜOĞLU, G. (1998), *Dinleme ve okuma öğretimi*. Eskişehir: Anadolu Universitesi Açık Öğretim Fakültesi Yayınları. No:587.

ÖZBAY, M. (2005). Bir dil becerisi olarak dinleme eğitimi. Ankara: Akçağ Yayıncılık.

SEVER, S. (2010). Türkçe öğretimi ve tam öğrenme. Ankara: Anı Yayıncılık.

TEMUR, T. (2010). Dinleme metinlerinden önce ve sonra soruların üniversite öğrencilerinin dinlediğini anlama becerisi düzeyine etkisi. *Selçuk Universitesi Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, 29, 305-320.

UNGAN, S. (2007). İlköğretimde Türkçe öğretimi (Editor: A. Kırkkılıç, H. Akyol), Ankara: Pegem Akademi Yayıncılık.