Arch Health Sci Res. 2024;11(3):148-154
DOI: 10.5152/ArcHealthSciRes.2024.23106

Original Article

Testing Validity and Reliability of Cross-linguistic Lexical Task—Turkish (CLT-TR) in Preschoolers with and without at Risk of Language Disorder

Nevin YILMAZ ÇİFTECİ[1], Aylin Müge TUNÇER[2]

[1]Department of Speech and Language Therapy, İzmir Bakırçay University, Faculty of Health Sciences, İzmir, Türkiye
[2]Department of Speech and Language Therapy, Muğla Sıtkı Koçman University, Faculty of Health Sciences, Muğla, Türkiye

This study is derived from the doctoral dissertation of the corresponding author in Eskişehir Anadolu University, Graduate School of Health Sciences, Speech and Language Therapy Doctorate Program.

Cite this article as: Yılmaz Çifteci N, Tunçer AM. Testing validity and reliability of cross-linguistic lexical task—Turkish (CLT-TR) in preschoolers with and without at risk of language disorder. *Arch Health Sci Res.* 2024;11(3):148-154.

ABSTRACT

Objective: Lexical knowledge provides important insights into language development of children. The assessment of this knowledge through valid and reliable tools is an essential step in diagnosing any linguistic problems. The recent study aims to test validity and reliability of the Cross-linguistic Lexical Task—Turkish (CLT-TR) among monolingual preschoolers with and without developmental language disorder. Cross-linguistic Lexical Task is developed to measure the lexical knowledge of mono/bi/multilingual preschool children. The test has 34 different language versions, including Turkish. Cross-linguistic Lexical Task—Turkish consists of comprehension and production subtests for both nouns and verbs with a total of 128 items.

Methods: The participants were 245 typically developed (TD) and 30 children at risk of developmental language disorders (r-DLD), all of whom were Turkish-speaking monolingual children aged between 2.00 and 4.11 years. CLT-TR administered individually to the participants. Maximum score was 30 for each subtest. Scores were compared within and between groups.

Results: The results showed that CLT-TR scores change significantly between developmental groups. Typically developed group had significantly higher scores than children in r-DLD group. Moreover, the test had construct validity evidence by showing significant differences in CLT-TR scores between age groups and parental education levels. These differences were also observed within subtests regarding receptive and productive language skills in nouns and verbs. According to the reliability analysis, the test has high internal consistency, stability, and objectivity.

Conclusion: According to the findings, it can be concluded that CLT-TR is a valid and reliable tool for measuring lexical knowledge of monolingual preschoolers. Cross-linguistic Lexical Task can successfully differentiate TD children from children with r-DLD. This test can be utilized by various professionals working with these age groups across different fields to gain valuable insights into language development firsthand.

Keywords: Lexical development, developmental language disorder, preschoolers, cross-linguistic, test validity, reliability

Introduction

Children's lexicon is an essential indicator of language emergence and development.[1] Lexical development requires the integration of phonological, semantic, and morpho-syntactic knowledge with cognitive and social processes. Moreover, it serves as a prerequisite for acquiring the more complex grammatical structures of a language.[2] Lexical development encompasses how children build their vocabulary, attribute meanings to words at different ages, and how these meanings can be altered through various experiences and contexts.[3]

Cross-cultural studies of typical lexical development reveal that children begin producing their first words around the age of 12 months. During the developmental process, children's vocabulary naturally expands with age.[4] As an early indicator of language emergence and risk of language disorders, assessment of lexical skills has a significant role in monitoring language progression in preschool years. Both typically developing (TD) children and children with developmental language disorder (DLD) comprehend far more words than they can produce.[4,5] In vocabulary tasks, children with DLD have lower scores in both receptive and expressive skills than TD children; however, limited expressive lexical knowledge is considered a more sensitive marker of DLD.[6]

Received: August 26, 2023
Revision Requested: December 20, 2023
Last Revision Received: March 26, 2024
Accepted: April 3, 2024
Publication Date: July 12, 2024

Corresponding author: Nevin YILMAZ ÇİFTECİ, e-mail: nevin.yilmaz@bakircay.edu.tr

Lexical skills of preschool children can be assessed by using different tools such as MacArthur-Bates Communication Development Inventory (CDI),[7] Language Development Survey (LDS),[8] and Peabody Picture Vocabulary Test V (PPVT-V).[9] These well-known tools are adapted for different languages and are widely used by professionals. The first 2 tests assess vocabulary skills based on parental reports. Although parental reports are cost-effective and easy to apply, they are not capable of reflecting children's vocabulary entirely.[1] Also, after the second year of life, children's vocabulary, especially the receptive one, develops so rapidly that parents may not be able to monitor it accurately anymore. The last test, PPVT-V,[9] assesses only receptive vocabulary. Besides, all these tools are developed for the assessment of monolinguals and cannot be used for bi/multilingual children.

Lexical development of both mono/bi/multilingual children has similar patterns.[10] However, when bi/multilinguals' lexical knowledge is assessed only in one language, they seem to have lower performance compared to the monolinguals.[11] Because of this similarity, unless the bi/multilingual children are assessed using appropriate tools, they may have a risk of misdiagnosis. Considering this problem, researchers in related fields came together in COST Action (IS0804) namely Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment (LITMUS) project, and developed assessment tools specifically used for bi/multilingual children.[12]

Cross-linguistic Lexical Task, as a part of LITMUS battery, is an assessment tool that is designed for assessing lexical knowledge of bi/multilinguals cross-linguistically in preschool years.[12] Cross-linguistic Lexical Task directly assesses lexical skills in both comprehension–production tasks and both word classes of nouns and verbs. Because CLT is administered directly to children, it presents a chance to observe children's productions, including pronunciation and lexical errors as well.

Cross-linguistic Lexical Task has been adapted for 34 different languages, including Turkish (TR). Various versions of CLT have been used in comparing lexical knowledge of monolinguals, bilinguals, and multilinguals[13-15] with TD and DLD.[16-18] Cross-linguistic Lexical Task has already been adapted to Turkish and used in comparing lexical skills of monolinguals with bilinguals, as well as TD with DLD children's lexical skills and age groups.[19] Cross-linguistic Lexical Task has been used in many different researches and presented well-documented results that CLT scores increase with age;[20] TDs have higher scores than DLDs,[17-21] and children have better performance on comprehension than production,[13-20-22] as well as in nouns than verbs.[13-15] Yet there is no available study on the validity and reliability of CLT in any language version.

As CLT is a standardized assessment tool, it is important to establish its validity, which refers to how accurately it measures the intended vocabulary skills. Cross-linguistic Lexical Task already had content validity because its tasks were created after long theoretical research, and items of the tasks were selected and determined through expert judgement.[12-23] As CLT is developed to assess the lexical knowledge of children with DLD and bilinguals, it is essential to provide how valid and reliable CLT is before use to prevent possible misdiagnosis.

Lexical knowledge of Turkish-speaking children may be assessed with vocabulary subtests as a part of more general language assessment tools like the Turkish Early Language Development Test (TEDİL).[24] However, these subtests have few items that provide limited information about children's lexical development. The Turkish version of the CDI is based on parental reports, and the Turkish Expressive and Receptive Language Test (TIFALDI)[25] has established validity and reliability only for the receptive language subtest; while PPVT-V,[9] which is standardized only for 3.01-3.11-year-old Turkish speakers, is used for receptive vocabulary skills.[26] There are limited valid and reliable lexical tools developed or adapted for Turkish-speaking children.

The current study aims to test the validity and reliability of the CLT-Turkish (CLT-TR) in monolingual preschoolers, both with and without the risk of DLD. St Clair[27] defines children at risk of DLD as those who have lower scores in a standardized language test than expected. Because of ongoing arguments related to the diagnosis and terminology used in language disorders, the term children at risk of DLD (r-DLD)[27,28] is preferred to be used because of the age groups included.

To the best of our knowledge, among the other versions of the CLT, CLT-TR is the first version to have studied the validity and reliability of the tool with such a large sample size. CLT-TR, which can directly assess lexical skills in the comprehension–production of nouns and verbs of preschool children, may help to meet the need for a valid and reliable lexical tool in Türkiye. This new tool can make an important contribution to the language assessment processes of the professionals in the related fields. Considering the population of the bi/multilingual children in Türkiye, it can be an important step to develop new tools that are designed to assess bi/multilingual children.

Methods

Participants

Participants of the study were 245 TD children (125 girls and 120 boys) and 30 children with r-DLD (13 girls and 17 boys) Turkish-speaking monolingual children aged between 2.00 and 4.11 years. The sample size was determined based on a power analysis conducted using G*Power version 3.1.9.7.[29] The results indicated that the required sample size to achieve 90% power for detecting a medium effect, at a significance criterion of $\alpha = 0.05$, was n = 166 ($n_{TD} = 138$, $n_{r\text{-}DLD} = 28$) for the Mann–Whitney *U*-test. Thus, the obtained sample size of n = 275 ($n_{TD} = 245$, $n_{r\text{-}DLD} = 30$) is adequate to test the study hypothesis.

The language skills of the participants were evaluated using a standardized test, as explained below. Children who scored average or above were included in the TD group, while children who scored below average, poor, and very poor were included in the children with r-DLD group. Participants in the r-DLD group were identified based on having lower linguistic skills in a standardized language test than expected.[27] In both developmental groups, children did not have any other developmental or health problems. The background information of participants is shown in Table 1.

Materials

TEDİL: TEDİL is the Turkish version of the Test of Early Language Development (TELD). The validity and reliability of TEDİL are

Table 1. Background Information About Participants

Age		n	%	M (SD)
TD	2.00-2.11	54	22.0	2;6 (3.2)
	3.00-3.11	91	37.2	3;5 (3.2)
	4.00-4.11	100	40.8	4;5 (3.5)
r-DLD	2.00-2.11	10	33.3	2;8 (1.6)
	3.00-3.11	11	36.7	3;5 (3.2)
	4.00-4.11	9	30.0	4;6 (5)

r-DLD, children at risk of developmental language disorder; TD, typically developing children.

established.[24] It is used to assess children's both receptive and expressive language skills aged between 2.00 and 7.11 years. TEDİL scores enable information about language skills as "Very Superior," "Superior," "About Average," "Average," "Below Average," "Poor," "Very Poor." This test was used to decide whether children have typical language ability or they are r-DLD.

CLT-TR: CLT is a picture-based task that consists of 4 subtests namely noun comprehension, verb comprehension, noun production, and verb production. Comprehension subtests include 1 target and 3 distractors. Children are expected to show the target picture out of 4 items. Production subtests, on the other hand, have only 1 picture in which children are asked to produce the target word. Each subtest includes 2 trials and 30 target words, and in total, there are 128 items. Development procedure and adaptation steps present content validity proof for CLT.[12] Turkish version of the task was developed[19] in 2013 and revised in 2022.[30] In revisions, some of the pictures were redrawn by the artist working for LITMUS-CLT project. The revised version was used in this study.

Procedure

Turkish Early Language Development Test and CLT-TR are implemented to the children individually and face-to-face in a quiet room at kindergartens. During CLT administration, production and comprehension tasks were counterbalanced across participants. For comprehension tasks, the experimenter asked, "Where is a (target noun)?" and "Which/Who is (target verb)?" and noted the number of the picture that children pointed to. For production tasks, the experimenter asked, "What is this?" for nouns and "What is he/she/it doing?/ What is happening?" for verbs. Verbal productions of the participants were recorded using an iPhone 7.

Ethical Considerations

Participants were sampled from kindergartens. Ethical approval was obtained from the Ethical Committee of the Anadolu University (Approval no: 68215 917-0 50.99 -E.43 658, Date: July 8, 2020) and the Ministry of National Education (Approval no: 49614 598-6 05.01-E.1812495 6). Written informed consent was obtained from parents of the children who participated in this study.

Statistical Analysis

The answers of the children were noted down on the scoring sheets. In comprehension tasks, the number of the picture, and in production tasks, the answers of the children were written as they pronounced it. When the answer is unclear to put under a certain category, an agreement was made through discussion with researchers. The maximum score for each subtest of the CLT-TR is 30 points. Data were examined in terms of Shapiro–Wilk test, kurtosis, and skewness to test normality of the dataset. Results showed that data were not distributed normally. Therefore, non-parametric analyses were conducted, and Bonferroni correction was used to adjust the significance level according to the number of comparisons made. Validity and reliability analyses include between-groups comparisons, correlational tests, consistency, stability, and objectivity analysis. All data were analyzed using the Statistical Package for Social Sciences version 23.0 software (IBM Corp.; Armonk, NY, USA).

Results

Validity of CLT-TR

Lexical skills of preschool children are expected to change over time, and scores on the lexical test are expected to vary in a predictable way as a function of membership in some group. So, we did some comparisons of CLT-TR scores between different developmental groups, age groups, and maternal education level groups.

For testing the construct validity of the CLT-TR, developmental groups' (TD and r-DLD) scores were compared by conducting Mann–Whitney*U*-test. The mean scores, standard deviations (SD), and between-groups comparison results are shown in Table 2.

As shown in Table 2, children with r-DLD had significantly lower scores than TD children in all subtests of CLT-TR. In other words, the CLT-TR scores present significant differences in lexical skills of TD and children with r-DLD groups.

Age

CLT-TR scores were compared between age groups to find more evidence for the construct validity of the test. Mean scores and SDs for TD age groups and Kruskal–Wallis test results are shown in Table 3.

As seen in Table 3, there was a significant difference between age groups in all subtests of CLT-TR. To find out which age groups differ significantly, Mann–Whitney *U*-test and Bonferroni correction were conducted. The results are presented in Table 4.

Results showed that the 2-year-old group had significantly lower scores than the 3 and 4-year-old groups; and the 3-year-olds had significantly lower scores than the 4-year-olds in all subtests of CLT-TR, as shown in Table 4.

Language Task and Word Class

The CLT-TR has 2 tasks, namely, comprehension and production, and includes 2 word classes: nouns and verbs. To find more evidence for construct validity of the CLT-TR, scores of 4 subtests in TD and children with r-DLD groups were compared with Friedman test. The results are shown in Table 5.

According to the results of the Friedman test presented in Table 5, significant differences within groups were observed in in CLT-TR subtests. To find out which subtests differ within groups, the Wilcoxon signed-rank test was conducted. This test was used for TD and children with r-DLD groups separately.

Table 2. Between Groups Comparison of CLT-TR Scores

Subtest	Group	n	Minimum	Maximum	Mean	SD	Mean Rank	Sum of Ranks	*U*	*P*
N-Comp	TD	245	9	30	25.54	4.56	146.55	35902.00	1583.0	<.001
	r-DLD	30	5	28	20.36	5.92	68.27	2048.00		
V-Comp	TD	245	5	30	22.06	5.70	145.80	35722.00	1763.0	<.001
	r-DLD	30	4	28	16.76	5.57	74.27	2228.00		
N-Prod	TD	245	0	30	19.41	5.83	148.56	36396.00	1089.0	<.001
	r-DLD	30	1	19	10.93	5.86	51.80	1554.00		
V-Prod	TD	245	0	28	16.22	6.09	147.00	36015.00	1470.0	<.001
	r-DLD	30	0	18	9.26	5.50	64.50	1935.00		

Comp, comprehension; N, noun; Prod, production; V, verb.

Table 3. CLT-TR Scores between Age Groups (n = 245)

Subtest	Age Groups	n	Minimum	Maximum	Mean	SD	Mean Rank	X^2	*df*	*P*
N-Comp	2.00-2.11	54	9	29	20.38	5.02	49.94			
	3.00-3.11	91	16	30	25.48	3.29	109.70	1115.06	2	<.001
	4.00-4.11	100	17	30	28.39	2.25	174.55			
V-Comp	2.00-2.11	54	5	29	15.40	5.34	47.85			
	3.00-3.11	91	9	30	21.86	4.21	112.37	113.53	2	<.001
	4.00-4.11	100	15	30	25.83	3.20	173.25			
N-Prod	2.00-2.11	54	0	24	12.88	5.66	51.74			
	3.00-3.11	91	9	30	19.19	4.37	113.54	100.72	2	<.001
	4.00-4.11	100	12	29	23.14	3.48	170.09			
V-Prod	2.00-2.11	54	0	24	8.87	5.10	44.19			
	3.00-3.11	91	7	27	15.73	3.88	109.75	129.39	2	<.001
	4.00-4.11	100	7	28	20.64	3.78	178.06			

Comp, comprehension; N, noun; Prod, production; V, verb.

Considering the language task, test results show that TD participants have significantly higher scores in comprehension than production, and this difference exists for both word classes.

When it comes to the word classes, the TD group had significantly higher scores in nouns than verbs in both comprehension and the production tasks.

Like the TD participants, for children with r-DLD group, there is also a significant difference between nouns and verbs and also comprehension and production tasks. Findings for TD children and children with r-DLD group are shown in Table 6.

Altogether, about lexical tasks, participants in both groups have significantly better performance in comprehension than production and are related to the word classes better in nouns than verbs.

Parental Education

Parental education groups were divided into 3 categories as "Elementary," "High school," and "University." All subtests and maternal/paternal education levels were compared separately. For maternal education level, Kruskall–Wallis test results showed that there were significant differences in scores of production tasks but not in the comprehension tasks ($P < .05$). Paired comparisons were done by Mann–Whitney *U*-test. According to the results, children whose maternal education level was university had higher scores in N-Prod subtest than elementary ($U = 1318.50$, $Z = -3.236$, $P < .001$) and high school levels ($U = 2940.00$, $Z = -2.608$, $P < .001$). Also, they had better performance than the high school level in V-Prod subtest ($U = 2992.00$, $Z = -2.456$, $P < .001$). No significant differences were found between the elementary and high school levels in the V-Prod.

For paternal education level comparisons, the same analysis was conducted, and it was found that the only significant difference is between elementary and university levels in the N-Prod subtest ($U = 1116.0$, $Z = -3.341$, $P < .001$). There is no significant difference between other educational levels and in any other subtests.

Table 4. Paired Comparison of Age Groups' CLT-TR Scores

Subtest	Age Groups	*U*	*Z*	*P*
N-Comp	2-3 years	956.50	−6.156	<.001
	2-4 years	255.50	−9.369	<.001
	3-4 years	1839.50	−7.198	<.001
V-Comp	2-3 years	844.00	−6.612	<.001
	2-4 years	255.00	−9.282	<.001
	3-4 years	1970.00	−6.787	<.001
N-Prod	2-3 years	976.00	−6.068	<.001
	2-4 years	333.00	−8.978	<.001
	3-4 years	2208.00	−6.155	<.001
V-Prod	2-3 years	673.00	−7.311	<.001
	2-4 years	228.50	−9.376	<.001
	3-4 years	1572.50	−7.823	<.001

Bonferroni correction: $P \leq .05/3 \rightarrow .017$.
Comp, comprehension; N, noun; Prod, production; V, verb.

Reliability of CLT-TR

In the scope of reliability analyses, consistency was tested by using Cronbach's alpha (CA), stability was analyzed by test–retest, and lastly, the objectivity of the test was analyzed by inter-rater reliability.

Internal Consistency

In the scope of reliability test, CA test was conducted, and subtest–total test score correlations for internal consistency and test–retest and inter-rater reliability analysis were calculated. The CA value of CLT-TR showed that the tool has good internal consistency (CA = 0.96).

To test the internal consistency of the CLT-TR, correlations between subtests of the CLT-TR and the subtest–total score of the test were calculated. Spearman correlation coefficients showed that each subtest of the CLT-TR has a significant positive correlation with the other subtests and with the total test score ($P < .01$).

Test–Retest Reliability

For test–retest reliability, 10% of TD participants were randomly selected and retested after a 2-3 weeks interval. The Spearman correlation coefficients were significant for all subtests, which were N-Comp ($r_s = 0.82$, $P < .001$), V-Comp ($r_s = 0.95$, $P < .001$), N-Prod ($r_s = 0.90$, P

Table 5. Within Group Comparison of CLT-TR Scores

Group	Subtest	n	Mean Rank	X^2	*df*	*P*
TD	N-Comp	245	3.89	876.921	4	<.001
	V-Comp		2.80			
	N-Prod		2.06			
	V-Prod		1.24			
r-DLD	N-Comp	30	3.73	98.200	4	<.001
	V-Comp		2.87			
	N-Prod		1.95			
	V-Prod		1.45			

Comp, comprehension; N, noun; Prod, production; V, verb.

Table 6. CLT-TR Scores' Comparison According to Lexical Task and Word Class in TD and r-DLD Groups

Task		Word Class	Ranks	n	Mean Rank	Sum of Ranks	*Z*	*P*
TD group	Comprehension	Verb	Negative	215	115.64	24862.00		
		Noun	Positive	10	56.30	563.00	−12.456	<.001
			Ties	20				
			Total	245				
	Production	Verb	Negative	183	121.41	22218.00		
		Noun	Positive	37	56.54	2092.00	−10.669	<.001
			Ties	25				
			Total	245				
	Word class	*Task*						
	Noun	Production	Negative	238	122.84	29236.50		
		Comprehension	Positive	4	41.63	166.50	−13.348	<.001
			Ties	3				
			Total	245				
	Verb	Production	Negative	233	123.33	28735.00		
		Comprehension	Positive	7	26.43	185.00	−13.274	<.001
			Ties	5				
			Total	245				
Task		**Word class**	**Ranks**	**n**	**Mean Rank**	**Sum of Ranks**	***Z***	***P***
r-DLD group	Comprehension	Verb	Negative	24	13.71	329.00		
		Noun	Positive	2	11.00	22.00	−3.911	<.001
			Ties	4				
			Total	30				
	Production	Verb	Negative	20	15.25	305.00		
		Noun	Positive	7	10.43	73.00	−2.807	.005
			Ties	3				
			Total	30				
	Word Class	*Task*						
	Noun	Production	Negative	28	15.93	446.00		
		Comprehension	Positive	2	9.50	19.00	−4.394	<.001
			Ties	0				
			Total	30				
	Verb	Production	Negative	26	15.15	394.00		
		Comprehension	Positive	2	6.00	12.00	−4.355	<.001
			Ties	2				
			Total	30				

Comp, comprehension; N, noun; Prod, production; V, verb.

< .001), and V-Prod (r_s = .85, P < .001). Results support that CLT-TR has high stability over time.

Inter-rater Reliability

Lastly, we tested inter-rater reliability by randomly selecting 30 children among all participants, and records of them were sent to a speech and language pathologist (SLP). Necessary background information about the study and the CLT-TR was given to the SLP. Spearman correlation coefficients were found significantly high for all subtests, which are N-Comp (r_s = 0.99, P < .001), V-Comp (r_s = 0.99, P < .001), N-Prod (r_s = 0.99, P < .001), and V-Prod (r_s = 0.96, P < .001), which shows CLT-TR scores consistent when different raters scoring the test.

Reliability analyses show that CLT-TR is a reliable tool to assess lexical skills by having good internal consistency, high stability, and good objectivity.

Discussion

The present study investigated the validity and reliability of CLT-TR in assessing monolingual preschoolers' lexical skills. In the scope of this aim, discriminative and construct validity analyses were conducted by comparing participants' scores between and within groups. Then to test the reliability of the CLT-TR, internal consistency, test–retest, and inter-rater reliability analyses were done.

Construct Validity

Construct validity is directly related to the theoretical and conceptual structure of the test.[31] In language development, more specifically in vocabulary development, the differences related to age, tasks, comprehension-production, and the word class, nouns–verbs, are a theoretically accepted distinction.[20] Moreover, children's linguistic performance can be affected by parameters of socioeconomic status (SES) like parental education.[32] To establish the construct validity of CLT-TR, efforts were made to gather evidence by comparing scores between developmental groups, age groups, language tasks, word classes, and parental education level.

To begin with, the comparison between developmental groups' results revealed that TD children have significantly higher scores than children with r-DLD in all subtests. Similar findings were reported from several studies using CLT.[17-18-21] Previous studies exhibited that DLD children have limited vocabulary and limited word-meaning organization skills.[33] In the early stages, DLD children have problems acquiring new words[6] which leads to limited vocabulary size compared to TD children. It is claimed that DLD children have some troubles related to short-term memory[6]; working verbal memory and verbal short-term

memory and maintenance of attention.[34] Cross-linguistic Lexical Task—Turkish includes both linguistic and executive functions skills processes such as eliminating distractors, keeping the target word in memory, and recalling the correct/target word. Limitations in linguistic and executive function skills may cause poorer performance in DLD children. Parallel with the literature findings, our results can provide proofs that CLT-TR significantly discriminates TD children from children with r-DLD in terms of vocabulary.

Age

In the scope of construct validity analysis, CLT-TR scores of participants in different age groups in TD children were compared. Findings showed that CLT-TR scores differed significantly between age groups, in which younger groups had lower scores than the older groups. It is a well-established finding that in preschool years vocabulary size increases with age.[5] Studies that used different versions of CLT also reported similar findings that CLT scores differed between age groups and also increased when the age of participants increased.[20-36] When CLT items were selected, the age of word acquisition was one of the parameters to include the target word in the task. It revealed that this parameter is a good predictor for lexical knowledge,[22] and lexical skills and age are positively correlated.[15-18]

Language Task and Word Class

Cross-linguistic Lexical Task—Turkish includes comprehension and production tasks for both word classes nouns and verbs. Subtest scores of TD children and children with r-DLD were compared within groups according to the task demanded and word class. Task-related comparisons showed that participants had significantly better performance in comprehension than production and are related to the word class they had significantly higher scores in nouns than verbs.

Findings supported the previous studies that exhibited preschoolers had lower scores in lexical production than comprehension. Studies using different versions of CLT also showed similar results.[13-16-20-22-34-36] In preschool years, comprehension skills develop before production skills do and provide valuable information about lexical and language development, also TD children understand far more words and simple expressions than they produce in the early years.[4]

Related to the word classes, literature findings revealed that nouns are more dominant than verbs in early years and among word classes, it is shown that the ratio is significantly higher in favor of nouns than the verbs in preschoolers' lexicon.[37] Studies with CLT obtained similar results as well in which children had better performance in nouns subtests than verbs.[13-15-19] This situation is observed for both TD children and children with r-DLD. Research suggested that DLD children put more effort into comprehending the meanings of verbs, learning new verbs more slowly,[6] and using verbs more limitedly than nouns.[12] It is claimed that verbs are "more difficult" than nouns because of several reasons.[37] Verbs have more complicated semantic and lexical organizations than nouns. While nouns share more common features and can be placed in a certain semantic category more easily, this situation is different for verbs. As verbs refer to more elements such as subject, actor, vehicle, and location which are always associated with a noun, when they are organized, they are subjected to more than one classification, unlike more simply organized concrete nouns. Lastly, in many languages, including Turkish, verbs have more inflections than nouns do. In lexical tasks, when verbs are the target word, children do not only have access to the verb itself but also to specific and context-related inflections that verbs possess.[37] Therefore, verb tasks turn into more complicated tasks than noun tasks.

Parental Education Level

Parental education as an important indicator of SES is positively correlated with lexical knowledge of children.[32] Participants' CLT-TR scores were compared between maternal and paternal education groups. Results showed that children who had higher maternal education levels had higher scores in production subtests, while paternal education only differed only in noun production subtests. Other studies with CLT found inconsistent results. While some of them[14-18] revealed no significant differences between groups, another found higher performance of children on the behalf of higher education level.[38] Mother's education level is a stronger marker of SES than father's, and parental education level has an indirect impact on children's vocabulary, in which parents from higher education can provide richer language input during interactions.[32] Generally, in Türkiye, mothers are the main caregivers of children than fathers. This can be one explanation of why maternal education level differs in more subtests than fathers did. Also, participants in the lower education group were fewer than the higher ones.

Reliability

Within the scope of internal consistency analysis, the CA value for the overall test was found to be 0.96, which is interpreted as "excellent" internal consistency. The second proof for internal consistency of CLT-TR was revealed by the correlation between the subtests with each other and with the total test score. Results showed there was a positive and significant correlation between all subtests and between subtest and total scores. It is stated that a correlation coefficient greater than 0.50 is evidence of a strong relationship between variables.[31] Current studies' findings, obtaining a value of more than 0.70, mean that there is a significant and "strong" correlation between the subtest–subtest and subtest–total test score. It was interpreted that the subtests are parallel with the CLT-TR's vocabulary measurement purpose and are necessary for this purpose.

Test–retest reliability measurements of the participants were found above 0.80 for all subtests. The correlation coefficient should be at least 0.70 to demonstrate that a scale is stable.[31] The results showed that the test was stable and had test–retest reliability.

Lastly, for objectivity analysis, the inter-rater reliability coefficient was calculated, and it was found to be greater than 0.90 for each subtest. This value is a coefficient interpreted as perfect.[31] In the audio recordings, the practitioner audibly expresses the number of picture the child has shown for the comprehension subtests. So, it is thought that it was inevitable to find high agreement between raters in the comprehension subtests. In the production subtests, the productions of the participants are written as they are heard. For this reason, it is thought that a high correlation coefficient in the production subtests is more important in terms of reliability to ensure inter-rater reliability.

Conclusion

To summarize, the research findings revealed that CLT-TR can discriminate developmental groups (TD and children with r-DLD) in terms of lexical knowledge. Cross-linguistic Lexical Task—Turkish has also construct validity evidence from age groups, language tasks (comprehension–production), word classes (noun–verb), and parental education level comparisons. Reliability findings showed that CLT-TR has high internal consistency, stability, and objectivity. As a main result of the study, it was revealed that CLT-TR is a valid and reliable tool that can be used to measure the lexical skills of monolingual preschoolers.

Ethics Committee Approval: Ethics committee approval was received for this study from the Ethics Committee of Anadolu University (Approval no: 68215 917-050.99-E.43658, Date: July 8, 2020).

Informed Consent: Written informed consent was obtained from parents of the children who participated in this study.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – N.Y.Ç., A.M.T.; Design – N.Y.Ç., A.M.T.; Supervision – A.M.T.; Materials – N.Y.Ç.; Data Collection and Processing – N.Y.Ç.; Analysis and Interpretation – N.Y.Ç., A.M.T.; Literature Search – N.Y.Ç., A.M.T.; Writing – N.Y.Ç.; Critical Review – A.M.T.

Acknowledgments: We thank Özlem Ünal-Logacev and Pınar Ege for their major contribution to the development of the first version of CLT-TR. We thank Ewa Haman and Magdalena Łuniewska to answer all our questions and encouragement.

Declaration of Interests: The authors have no conflict of interest to declare.

Funding: The authors declared that this study has received no financial support.

References

1. Bello A, Giannantoni P, Pettenati P, Stefanini S, Caselli MC. Assessing lexicon: validation and developmental data of the Picture Naming Game (PiNG), a new picture naming task for toddlers. *Int J Lang Commun Disord*. 2012;47(5):589-602. [CrossRef]
2. Dockrell JE, Messer DA. Lexical acquisition in the early school years. In: Berman RA, ed. *Language Development across Childhood and Adolescence-Trends in Language Acquisition Research Series (3)*. Amsterdam, Netherlands: John Benjamins Publishing Company; 2004:35-52.
3. Merriman EW. Lexical development. In: Brooks PJ, Kempe V, eds. *Encyclopedia of Language Development*. SAGE Publications, CA; 2014:344-350.
4. Leonard LB. *Children with Specific Language Impairment*. MIT Press, Cambridge; 2014:51-53.
5. Roseberry-McKibbin C, Hegde MN. *An advanced review of speech-language pathology: preparation for Praxis and comprehensive examination*. PRO-ED, Austin, TX: 2016:109-124.
6. McGregor KK, Van Horne AO, Curran M, Cook SW, Cole R. The challenge of rich vocabulary instruction for children with developmental language disorder. *Lang Speech Hear Serv Sch*. 2021;52(2):467-484. [CrossRef]
7. Fenson L, Dale P, Reznick JS, et al. *The Mac Arthur Communicative Development Inventories: User's Guide and Technical Manual*. San Diego: Singular Publishing Group; 1994.
8. Rescorla L. The Language Development Survey: A screening tool for delayed language in toddlers. *J Speech Hear Disord*. 1989;54(4):587-599. [CrossRef]
9. Dunn DM. *Peabody Picture Vocabulary Test–Fifth Edition (PPVT-5)*. Minneapolis, MN; 2018.
10. Conboy BT, Thal DJ. Ties between the lexicon and grammar: cross-sectional and longitudinal studies of bilingual toddlers. *Child Dev*. 2006;77(3):712-735. [CrossRef]
11. Simonsen HG, Haman E. LITMUS-CLT: A new way to assess bilingual lexicons. *Clin Linguist Phon*. 2017;31(11-12):811-817. [CrossRef]
12. Haman E, Łuniewska M, Pomiechowska B. Designing Cross-linguistic Lexical Tasks (CLTs) for bilingual preschool children. In: Armon-Lotem S, de Jong J, Meir N, eds. *Methods for Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*. Bristol: Multilingual Matters; 2015:196-240.
13. Altman C, Goldstein T, Armon-Lotem S. Quantitative and qualitative differences in the lexical knowledge of monolingual and bilingual children on the LITMUS-CLT task. *Clin Linguist Phon*. 2017;31(11-12):931-954. [CrossRef]
14. Bohnacker U, Lindgren J, Öztekin B. Turkish-and German-speaking bilingual 4-to-6-year-olds living in Sweden: Effects of age. SES and home language input on vocabulary production. *J Lang Teach Res*. 2016;1:17-41. http://hdl.handle.net/10092/12907.
15. Lindgren A. *Cross Linguistic Lexical Tasks (CLTs): En: kvalitativ analys av ordproduktion hos enspråkiga svenska och flerspråkiga svensk-turkiska 4; 0-6; 11- åringar.* [Master thesis]. Uppsala: Uppsala University; 2015.
16. Kapalková S, Slančová D. The vocabulary profile of Slovak children with primary language impairment compared to typically developing Slovak children measured by LITMUS-CLT. *Clin Linguist Phon*. 2017;31(11-12):893-909. [CrossRef]
17. Khoury Aouad Saliby C, Dos Santos C, Kouba Hreich E, Messarra C. Assessing Lebanese bilingual children: the use of cross-linguistic lexical tasks in Lebanese Arabic. *Clin Linguist Phon*. 2017;31(11-12):874-892. [CrossRef]
18. Öberg L. *Words and Non-words. Vocabulary and Phonological Working Memory in Arabic-Swedish-Speaking 4-7-Year-Olds with and without a Diagnosis of Developmental Language Disorder* [Dissertation]. Uppsala: Acta Universitatis Upsaliensis. Studia Linguistica Upsaliensia; 2020.
19. Ünal Ö, Tunçer AM, Ege P, Sözcük Bilgisi Alt Çalışma Grubu: Diller Arası Sözcük Bilgisi Testi: Türkçe. In: Topbaş S, Mavis İ, Yavuz H, et al., eds. *TÖDİL Özgül Dil Bozukluğu Projesi: Anadili Türkçe Olan Tekdilli ve İkidilli Okul Öncesi ve İlköğretim Çağı (2:0-9:0 Yaş) Çocuklarında Özgül Dil Bozukluğu Ölçme ve Değerlendirme Çalışması*. Ankara; 2013:122-129. https://search.trdizin.gov.tr/yayin/detay/612552/todil-ozgul-dil-bozuklugu-turkiye-projesi-anadili-turkce-olan-tekdilli-ve-ikidilli-okuloncesi-ve-ilkogretim-cagi20-90-yas-cocuklarinda-ozgul-dil-bozuklugunu-olcme-ve-degerlendirme-calismasi.
20. Haman E, Łuniewska M, Hansen P, et al. Noun and verb knowledge in monolingual preschool children across 17 languages: data from cross-linguistic lexical tasks (LITMUSCLT). *Clin Linguist Phon*. 2017;31(11-12):818-843. [CrossRef]
21. Öztekin B. *Typical and Atypical Language Development in Turkish-Swedish Bilingual Children Aged 4-7*. [Dissertation]. Uppsala: Acta Universitatis Upsaliensis Studia Linguistica Upsaliensia; 2019.
22. Hansen P, Simonsen HG, Łuniewska M, Haman E. Validating the psycholinguistic aspects of LITMUS-CLT: evidence from Polish and Norwegian. *Clin Linguist Phon*. 2017;31(11-12):910-930. [CrossRef]
23. Łuniewska M, Haman E, Hansen P. Is there a road to universal assessment of lexical knowledge in multilingual children? Cross-cultural aspects of Cross-linguistic Lexical Tasks. *Helt Fabelaktig*. 2016. pp. 147-165.
24. Topbaş S, Güven S. *Türkçe Erken Dil Gelişimi Testi (TEDİL)*. Ankara: Detay Yayıncılık; 2011.
25. Kazak Berument S, Güven AG. Türkçe ifade edici ve alıcı dil (TİFALDİ) testi: I. Alıcı dil kelime alt testi standardizasyon ve Güvenilirlik Geçerlik çalışması. *Turk Psikiyatri Derg*. 2013;24(3):192-201. https://dergipark.org.tr/en/pub/egeefd/issue/4902/67205.
26. Özekes M. Peabody Resim Kelime Testi 3;01-3;12 Yaş aralığı İzmir bölgesi standardizasyonu çalışması. *Ege Eğitim Derg*. 2013;14(1):90-107.
27. St Clair MC, Forrest CL, Yew SGK, Gibson JL. Early risk factors and emotional difficulties in children at risk of developmental language disorder: A population cohort study. *J Speech Lang Hear Res*. 2019;62(8):2750-2771. [CrossRef]
28. Toseeb U, St Clair MC. Trajectories of prosociality from early to middle childhood in children at risk of developmental language disorder. *J Commun Disord*. 2020;85:105984. [CrossRef]
29. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191. [CrossRef]
30. Yılmaz-Çifteci N, Tunçer M. Diller arası sözcük bilgisi testi Türkçe (CLT-TR): Geçerlik ve Güvenirlik ön çalışması. *Turk Klin J Health Sci*. 2022;7(3):739-745. [CrossRef]
31. Orlikoff RF, Schiavetti N, Metz DE. *Research Design in Communication Disorders. Evaluating Research in Communication Disorders. The Pearson Communication Sciences and Disorders Series*. London: Pearson; 2015: 100-150.
32. Fenson L, Marchman V, Thal D, Dale PS, Reznick JS, Bates E. MacArthur–Bates. *Communicative Development Inventories: Users Guide and Technical Manual*. Baltimore, MD: Paul H Brookes: 2007.
33. Sheng L, McGregor KK. Lexical-semantic organization in. Children with specific language impairment. *J Speech Lang Hear Disord*. 2010;53(1):146-159. [CrossRef]
34. Archibald LM, Gathercole SE. Short-term and working memory in specific language impairment. *Int J Lang Commun Disord*. 2006;41(6):675-693. [CrossRef]
35. Abbot-Smith K, Morawska-Patera P, Łuniewska M, Spruce M, Haman E. Using parental questionnaires to investigate the heritage language proficiency of bilingual children. *Child Lang Teach Ther*. 2018;34(2):155-170. [CrossRef]
36. Van Wonderen E, Unsworth S. Testing the validity of the Cross-Linguistic Lexical Tasks as a measure of language proficiency in bilingual children. *J Child Lang*. 2020:1-25. [CrossRef]
37. Masterson J, Druks J, Gallienne D. Object and action picture naming in three and five-year-old children. *J Child Lang*. 2008;35(2):373-402. [CrossRef]
38. Perold Potgieter A, Southwood F. A comparison of proficiency levels in 4- year-old monolingual and trilingual speakers of Afrikaans, isiXhosa and South African English across SES boundaries, using LITMUS-CLT. *Clin Linguist Phon*. 2016;30(2):87-100. [CrossRef]