# SENTENCE COMPREHENSION IN CHILDREN WITH DEVELOPMENTAL LANGUAGE DISORDERS

**Barbora Červenková[1], & Gabriela Solná[2]**
[1]*Faculty of Education, Palacký University/Faculty Hospital Brno (Czech Republic)*
[2]*Faculty of Medicine, Charles university/Hospital Agel, Ostrava – Vítkovice (Czech Republic)*

## Abstract

*Purpose:* Language comprehension is vital to social and educational development. Difficulties in sentence comprehension can be detected in most children with developmental language disorders (DLD) using age-appropriate instruments with high sensitivity. There are two main theoretical frameworks for the characterization of difficulties in language acquisition in children with (DLD): linguistic-based and cognitive-based theories. The aim of this study is a comparison of language comprehension in children with developmental language disorder (DLD) and children with typical language development (TLD). Two types of tests referred to as main theoretical frameworks have been used (Token test and TEPO - Sentence Comprehension Test). *Sample:* We studied 101 children aged 4,3-6,9 years with (DLD) and typically developing peers (TLD). Data collection took place in the period 2021-2023. *Method:* It is quantitative research. A total of 47 children with (DLD) were compared with 54 children with (TLD) in two comprehension tests. The research goals were set as follows: 1) Correlation of TEPO test and Token test using Spearman´s correlation coefficient; 2) Comparison of homogeneity of test results in TLD and DLD children using Stuart–Maxwell test; 3) Determination of the sensitivity and specificity of both tests. *Results:* There is a high dependence between the two tests, indicating a high level of correlation, with $r_s$ = 0.81 in the whole research sample. There was a statistically significant difference ($p < 0.001$) between the test results based on percentile rank scores. The non-parametric unpaired Stuart-Maxwell test was used. The TEPO test has high sensitivity (95,7%) and the Token test has high specificity (90,7%). *Discussion:* Both comprehension tests have several advantages in terms of their characteristics, and they also have some limitations. The research results provide more information about the differences between the two testing methods and explain how the test results can be used in education.

***Keywords:*** *Specific language impairment, comprehension, language disorders.*

## 1. Introduction

Developmental language disorder (DLD) is a neurodevelopmental disorder with a prevalence rate of approximately 7% in early childhood (Siu, 2015). DLD manifests as difficulties in expressive language or both expressive and receptive language (American Psychiatric Association [DSM-V], 2013; Tomblin et al., 1997). DLD is a heterogeneous category that has consequences in the linguistic domains: phonology, word finding and semantics, syntax, pragmatics, and discourse (Kamhi & Clark, 2013) and also in cognitive domains: selective attention, attentional shifting, working memory and verbal short term memory (Archibald & Gathercole, 2006; Lum, Conti-Ramsden, Page, & Ulman, 2012; Lum & Zarafa, 2010; Willinger et al., 2017) and is often associated to general processing limitations (e.g., Kamhi & Clark, 2013). DLD persists from childhood to adolescence and adulthood (Kamhi & Clark, 2013). Problems with language development can cause significant interference with everyday life or educational progress (Bishop et al., 2017).

There are a variety of standardized tests available to assess sentence comprehension in children with Developmental Language Disorder (DLD). Some commonly used tests include: The Test for Reception of Grammar, The Clinical Evaluation of Language Fundamentals, The Comprehensive Assessment of Spoken Language, The Peabody Picture Vocabulary Test.

The Test for Reception of Grammar – TROG-2 (Bishop, 2003) measures a child's ability to understand and use grammatical structures in sentences. It includes a series of picture stimuli that the child is asked to describe using grammatically correct sentences. The Clinical Evaluation of Language Fundamentals – CELF -5 (Semel, Wiig, & Secord, 2013) measures a range of language skills, including sentence comprehension. The test includes a variety of tasks that assess a child's ability to understand

spoken sentences, including sentences with complex syntax and semantics. The Comprehensive Assessment of Spoken Language – CASL2 (Carrow-Woolfolk, 2017) assesses a range of language skills, including sentence comprehension. It includes a variety of tasks, such as sentence completion and comprehension questions, to assess a child's understanding of complex sentence structures. The Peabody Picture Vocabulary Test – PPVT-4 (Dunn, 2018) while not specifically designed to assess sentence comprehension, this test measures a child's ability to understand and use vocabulary words in sentences, which is an important component of overall language comprehension.

It is important to note that standardized tests are just one tool used to assess a child's language abilities, and they should be used in conjunction with other assessments, clinical observations, and reports from parents and teachers to fully understand the child's language profile. A qualified speech-language pathologist can help to determine the most appropriate assessments for a child with DLD and interpret the results in the context of the child's individual strengths and weaknesses.

## 2. Design

In the research methods of data collection, the oral speech understanding component was examined using the Token Test – TT (Bolceková, Preiss, & Krejčová, 2015) and the Sentence Comprehension test –TEPO (Solná & Červenková, 2022). Sentence comprehension tests were chosen for our research, rather than tests assessing comprehension at the level of single words, because tests based on sentence comprehension place high demands on short-term memory load due to the need to process and interpret grammatical relations, hence more complex results are obtained for children with DLD.

Our aim is to compare two tests of sentence comprehension based on different theoretical frameworks. The Token test is based on Cognitive theory. The children's performance in this test relies on short-term/working memory, attentional shifting and attentional control, inhibition and cognitive flexibility and planning (Kamhi & Clark, 2013; Willinger et al., 2017; Schmoeger et al., 2019). Although the Token test (De Renzi & Vignolo, 1962) was not originally designed to test comprehension in children, later was shown to be efficacious in the detection of language problems in children with DLD (e.g., Cole & Fewell, 1983), and in the children´s version Token Test for Children –TTFC-2 (McGhee, Ehrler, & DiSimoni, 2007) is a widely used diagnostic method. In our research, we used a shortened version of TT (Bolceková, Preiss, & Krejčová, 2015). There are 35 items in total, divided into six sections. In this test, the child has to manipulate with plastic objects (tokens) which differ in size, form, and color.

Sentence comprehension test – TEPO (Solná & Červenková, 2022) is similarly to TROG 2, based on Linguistic theory and is intended for children from 3 to 8 years of age for assessing the level of language skills. The test has a multiple-choice format with lexical and grammatical foils. The child listens to a spoken sentence and must select one of four pictures to match what is heard. The items are organised into 15 blocks of 4 items, with the grammatical complexity of the blocks increasing as the test progresses. The standardized version included norms for 863 children. The administration time is 10 minutes.

## 3. Objectives

The research questions were set as follows: 1) What is the correlation of TEPO and TT using Spearman´s correlation coefficient; 2) What is the homogeneity of the results of both tests in a group of children with TLD and DLD using Stuart–Maxwell test; 3) What is the sensitivity and specificity of TEPO and TT in a subgroup of children with DLD.

## 4. Methods

### 4.1. Participants

Nineteen girls (40%) and twenty-eight boys (60%) aged 4-6 years (M = 5.9 years) were recruited at the Hospital AGEL, Ostrava-Vítkovice. DLD was diagnosed by a speech and language therapist whereby diagnosis implied at least two below average scores out of five in standardized language tests. Furthermore, a control sample of typically developing, age-matched children, 26 boys, and 28 girls, were invited to participate from kindergarten through Grade 1 in a public school in Ostrava. Children in the control group were tested with the same set of tests as the children with DLD. They were included in this group if they achieved normal test results. The characteristics of the children are shown in Table 1.

*Table 1. Characteristics of the children with developmental language disorder (DLD), and normally developing children with typical language development (TLD).*

| Characteristics (n = 101) | TLD (n = 54) | DLD (n = 47) | p-value |
|---|---|---|---|
| girls | 26 (58%) | 19 (42%) | 0,436* |
| boys | 28 (50%) | 28 (50%) | |
| girls age in months (median, IQR) | 67 (61-76)*** | 72 (66-83) *** | 0,103** |
| boys age in months (median, IQR) | 66 (61-71) *** | 69,5 (66-76) *** | 0,079** |

*chí-square test, **Mann-Whitney test, ***IQR – interquartile rank (25.-75. percentile)

## 4.2. Statistic analysis

Basic descriptive statistics (median, IQR – interquartile range: 25th-75th percentile, absolute counts, %) were used for the data analysis. The Chi-square test and symmetry test for qualitative traits were used for statistical evaluation. For quantitative traits, normality was tested using the Shapiro-Wilk test. Normality was confounded, children with DLD were older than those with TLD, p=0.020, so the non-parametric Mann-Whitney test was further used for statistical testing. The dependence between the TEPO and Token test results was assessed using ordinal Spearman's correlation coefficient. Sensitivity, specificity, and predictive values were computed. Statistical tests were evaluated at the 5% significance level. Stata version 16 software was used for processing.

## 4.3. Correlation of TEPO and TT using Spearman correlation coefficient

The correlation of the TEPO test with the Token test using Spearman's rank correlation coefficient in the entire participant population is $r_s$ = 0.81, in children without diagnosis (TLD) is $r_s$ = 0.63, and in children with DLD is $r_s$ = 0.58. Thus, the correlation of the TEPO Test with the Token Test according to Spearman's correlation coefficient according to the interpretation of De Vaus (2002) is moderate to substantial (0.50 - 0.69) in subgroups and substantial to very strong (0,70-0,89) in the whole research sample, all results are statistically significant.

## 4.4. Comparison of homogeneity of test results in TLD and DLD children

The distribution of children's scores in percentile ranks expressed as a percentage can be seen in Table 2.

*Table 2. Comparison of TEPO and TT results by percentile ranks in the whole research sample.*

| TT | TEPO | | | |
|---|---|---|---|---|
| | Norm | 5.perc. | 10.perc | Total |
| Norm | 43 (42,57%) | 12 (11,88%) | 10 (9,90%) | 65 (64,36%) |
| 5.perc. | 2 (1,98%) | 23 (22,77%) | 0 (0%) | 25 (24,75%) |
| 10.perc | 3 (2,97%) | 7 (6,93%) | 1 (0,99%) | 11 (10,89%) |
| Total | 48 (47,52%) | 42 (41,58%) | 11(10,89%) | 101 (100%) |

48% of children were within a normal limit in TEPO and 64 % of children in TT. 11% of children were at the 10th percentile in TEPO and 42% of children were at the 5th percentile. 11% of children were at the 10th percentile in TT and 25% were at the 5th percentile. There was a statistically significant difference in the symmetry test (p < 0.001) between the test results based on percentile rankings.

### 4.4.1. Comparison of homogeneity of test results by subgroups.
Comparison of TT and TEPO test results in TLD children: there was no statistically significant difference between the test results based on percentile rankings. For this evaluation, the Stuart-Maxwell test p=0.273 was used to assess the homogeneity of the quantitative variables. Comparison of TT and TEPO test results in DLD children: there was a statistically significant difference (p < 0.001) between the test results based on percentile rankings. For this evaluation the Stuart-Maxwell test p=0.0002 was used to assess the homogeneity of the quantitative variables.

## 4.5. Sensitivity and specificity of TT and TEPO in children with DLD (n=47)

Children with DLI in TEPO: 45 children are true positives and 2 children are false negatives. Children with DLI in TT: 31 children who are true positives and 16 children who are false negatives.

TEPO: the sensitivity of the test is 95.7%, the specificity of the test is 85.2%, the positive predictive value is 84.9%, and the negative predictive value is 95.8%.

TT: the sensitivity of the test is 66%, the specificity of the test is 90,7%, the positive predictive value is 86%, and the negative predictive value is 75,4%.

## 5. Discussion

The necessity for the evaluation of sentence comprehension in preschool and school-age children with DLD using short, effective, and feasible tools is obvious. For this reason, it is recommended by the American Speech-Language-Hearing Association or the American Academy of Pediatrics (Hagan, Shaw, & Duncan, 2008; Siu, 2015) early detection of language impairments. In this research two comprehension tests were compared: TT and TEPO. Both of these tests are short enough to be used for DLD screening in preschool age.

The correlation of the TEPO Test with the Token Test according to Spearman's correlation coefficient is very strong. If children develop typically and do not have comprehension difficulties, they can be tested with both tests with similar results. However, children with DLD score significantly differently on these tests. It is therefore necessary to define what the tests actually assess.

In our study, the sensitivity of TT in the DLD subgroup is 66%, and the specificity is 90,7%. We can compare these results with the study results of a team of authors (Willinger et al., 2017). In this study, the sensitivity of TT in the DLD subgroup is 69%. This study measured the psychometric properties of a short version of the Token test (containing 50 items) as a screening tool for detecting children with DLD in preschool age (4-6 years). Because the total score of the TTFC-2 test is lower than the recommended total classification rate of 80%, the authors of this study do not recommend using this test to detect comprehension disorders in preschool children under the age of 6. According to this group of authors, this test primarily measures children's intellectual capacity because the scores of children with DLD on the TTFC-2 test correlate significantly with poorer scores on the Wechsler Preschool and Primary Scale of Intelligence -WPPSI (Wechsler, 1967), specifically on the arithmetic and verbal subtest.

A later study (Schmoeger et al., 2019) showed that in age groups 4 and 5 years, Part V of the Token Test yielded acceptable classification rates (85.1% and 80.6%) whereas the age group 6 years was not significantly discriminated by any Token Test variable. Based on these results, it is clear that TT does not cover important linguistic aspects for identifying a heterogeneous group of patients, such as DLD, in several age groups.

TEPO targets specific linguistic features and therefore has 95,7% sensitivity and 85,2% specificity. Therefore, this test can be considered a suitable screening method for DLD. The TT can be used as the second test in the sequence, knowing that it evaluates a more general developmental factor in preschool children rather than an exclusive indicator for language comprehension.

## 6. Conclusion

According to the prevalence in which DLD occurs in the population, this condition can affect about two or three children out of every classroom, which may manifest itself in difficulties in comprehension. These problems are often not obvious either for parents or teachers. They can manifest themselves in children by: misunderstanding instructions, needing simpler language than peers, and asking for multiple repetitions. If the teacher notices these difficulties, he/she should recommend an assessment by a speech and language therapist. If the comprehension disorder is confirmed, the following strategies may be considered appropriate: creating an environment conducive to communication by reducing background noise, adapting to teachers' language, attracting the child´s attention before assigning a task, giving simple instructions in the order to be followed, avoid more abstract words, explaining abstract concepts, using familiar vocabulary, giving extra thinking time and processing time, inclusion pause between explaining key points and ideas, using gestures and repeating key points and summarizing of what has been said.

It is also important to encourage children to ask questions when they do not understand a word during lessons and then teach the meaning and practice it together. Encourage children to keep a record of new words that are introduced in a vocabulary book, a word wall display, or a word bank.

*References*

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.

American Speech-Language-Hearing Association. (n.d.). *Definitions of communication disorders and variations.* Retrieved May 22, 2023, from http://www.asha.org/policy/RP1993-00208/

Archibald, L. M., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders*, *41*(6), 675–693. doi:10.1080/13682820500442602

Bishop, D. (2003) *The Test for Reception of Grammar version 2 (TROG-2).* London: Psychological Corporation.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2017). Phase 2 of catalise: A multinational and multidisciplinary Delphi consensus study of problems with Language Development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080. doi:10.1111/jcpp.12721

Bolceková, E., Preiss, M., & Krejčová, L. (2015). *Token test pro děti a dospělé*. [Token test for children] Otrokovice: Propsyco.

Carrow-Woolfolk, E. (2017). *Comprehensive Assessment of Spoken Language, CASL-2* (2nd ed.) Torrance, CA: Western Psychological Services.

Cole, K. N., & Fewell, R. R. (1983). A quick language screening test for young children: The Token Test. *Journal of Psychoeducational Assessment, 1*(2), 149-153.

De Renzi, A., & Vignolo, L. A. (1962). Token test: A sensitive test to detect receptive disturbances in APHASICS. *Brain*, *85*(4), 665–678. doi:10.1093/brain/85.4.665

De Vaus, D. A. (2002). *Analyzing social science data*. Thousand Oaks, Calif: Sage.

Dunn, D. M. (2018). *Peabody Picture Vocabulary Test* (5th ed.). San Antonio, TX: Pearson Assessment.

Hagan, J., Jr., Shaw, J., & Duncan, P. (2008). *Bright futures: Guidelines for health supervision of infants, children, and adolescents* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics.

Kamhi, A. G., & Clark, M. K. (2013). Specific language impairment. In Dulac, O., Lassonde, M., & Sarnat, H. B. (Eds.), Handbook of clinical neurology (Vol. 111, pp. 219–227). Amsterdam: Elsevier.

Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, *48*(9), 1138–1154. doi:10.1016/j.cortex.2011.06.001

Lum, J. A., & Zarafa, M. (2010). Relationship between verbal working memory and the scan–C in children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *41*(4), 521–530. doi:10.1044/0161-1461(2010/09-0035

McGhee, R. L., Ehrler, D. J., & DiSimoni, F. (2007). *Token Test for Children* (2nd ed.) Austin, TX: Pro-Ed.

Semel, E., Wiig, E., & Secord, W. (2013). *Clinical Evaluation of Language Fundamentals*. The Psychological Corporation. San Antonio, TX: Pearson Clinical Assessment.

Schmoeger, M., Deckert, M., Eisenwort, B., Loader, B., Hofmair, A., Auff, E., & Willlinger, U. (2019). Evaluating Part V of the German version of the Token Test as a screening of specific language impairment in preschoolers. *Applied Psycholinguistics*. *41*(1), 237-258. doi.org/10.1017/S0142716419000493

Siu, A. L., & US Preventive Services Task Force. (2015). Screening for speech and language delay and disorders in children aged 5 years or younger: *US Preventive Services Task Force Statement. Pediatrics, 136*(2), e474–e481. https://doi.org/10.1542/peds.2015-1711

Solná, G., & Červenková, B. (2022). *TEPO-Test porozumění větám*. [TEPO - Sentence comprehension test]. Praha: Publi.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260. doi:10.1044/jslhr.4006.1245

Wechsler, D. (1967). *WPPSI Manual: Wechsler Preschool and Primary Scale of Intelligence*. New York: The Psychological Corporation.

Willinger, U., Schmoeger, M., Deckert, M., Eisenwort, B., Loader, B., Hofmair, A., & Auff, E. (2017). Screening for specific language impairment in preschool children: Evaluating a screening procedure including the token test. *Journal of Psycholinguistic Research*, *46*(5), 1237–1247. doi:10.1007/s10936-017-9493-z