

## Assessing the German vocabulary size of elementary school students in Graubünden, Switzerland

**Dominique Caglia**, Universität Freiburg (CH), Pädagogische Hochschule Graubünden  
**Erwin Tschirner**, Universität Leipzig, Herder Institut

*Vocabulary size correlates strongly with general language proficiency and is highly predictive of academic skills. Frequency-based vocabulary tests, such as the German VST (Vocabulary Size Test; Institut für Testforschung und Testentwicklung, 2019; Nation, 2001), are particularly informative. The present study validated this test with 229 fifth and sixth graders from Romansh- and German-instructed primary schools in Graubünden, Switzerland. The study shows a high reliability and internal validity of the German VST as well as a strong correlation of the test with the PPVT-4 (Lenhard et al., 2015) in a sub-sample. The test therefore provides relevant data on the academic vocabulary of students with German as L1, L2 or L3 in the Romansh context.*

### 1. Introduction

Students at Romansh schools in Graubünden, Switzerland, show heterogeneous German skills: While L1 Romansh students show similar German reading competences to L1 German students, students with L1 other than Romansh or German perform significantly worse in the same tests than their classmates (Peyer et al., 2014). Because German is the majority language of the canton, and the majority school language from secondary level, having sufficient German competences is crucial in the Romansh context. To determine valid and reliable measurements, standardized tests are needed to better classify and specifically develop the students' skills.

As vocabulary size correlates strongly with general language proficiency and is highly predictive of academic skills, we consider the assessment of students' general academic vocabulary to be decisive. One of the most common tests used to assess the vocabulary size of school children is the Peabody Picture Vocabulary Test (PPVT). The German version was normed in 2013 on a representative sample of 3,555 elementary and secondary school students in Germany (Lenhard & Lenhard, 2021). While it was shown that the PPVT is an economic and reliable assessment instrument, being based on pictures, i.e., on words that may be depicted visually, raises the question of whether it covers the vocabulary needed to be successful in educational settings. Another approach to defining vocabulary size is based on frequency, generally substantiated by the analysis of large written and spoken corpora. Some of the most common vocabulary size tests based on frequency are those developed by Nation (2001), or modeled on them. The ITT Vocabulary Size Test (VST) for German as a Foreign Language has been used in educational contexts since 2014, mostly with college students, and is highly predictive of L2 reading proficiency in academic contexts. The goal of this study was to examine whether it may also be used in elementary school contexts, and whether it provides meaningful results for L1 as well as L2 and L3 German in the context of Graubünden. The VST was administered to a total of 229 fifth and sixth grade students in Romansh- or German-instructed elementary schools. In addition, the PPVT (Version 4) was administered to a subset of 17 students. This article focuses on the reliability and internal validity of the VST, and whether it provides meaningful results for students with L1 German, Romansh, and other language backgrounds in elementary schools in Graubünden. The relationship between the VST and PPVT-4 is also explored.

### 2. Vocabulary and Reading

Receptive vocabulary size correlates highly with reading and listening proficiency and provides a measure of overall language proficiency (Schmitt, 2008; Stæhr, 2008; Milton, 2009). A study done by Hacking et al. (2019), for example, documented very high correlations between reading proficiency and receptive vocabulary size (Chinese:  $r = 0.84$ ; Russian:  $r = 0.87$ ; Spanish:  $r = 0.88$ ), while Stæhr (2008) found high correlations

between a receptive vocabulary size test and English reading ( $r = 0.83$ ), listening ( $r = 0.69$ ), and writing proficiency ( $r = 0.73$ ).

For general reading purposes, a vocabulary size of at least the 5,000 most frequent words of a language appears to be necessary. Readers need to comprehend between 95% and 98% of the tokens of a text to understand the text. Readers who understood 95% of the tokens of a text were able to grasp ca. 60% of the information in a text, while readers who comprehended 98% of the tokens acquired 70% of the information (Schmitt et al., 2011). Tschirner (2009) established that the 3,000 most frequent German words cover approx. 90% of the tokens in bestseller novels and 85% in newspapers, while the 5,000 most frequent words account for 93% of the tokens in novels and 88% in newspapers. A vocabulary size of 2,000 words is generally associated with a reading proficiency of A2, while vocabulary sizes of 3,000 and 5,000 words are associated with B1 and C1 respectively (Milton, 2010; Huhta et al., 2011; Tschirner, 2019). The B1 word list for German, for example, comprises some 3,400 lexemes (Glaboniat et al., 2016).

Due to the pragmatic-functional orientation of the Common European Framework of Reference (CEFR), CEFR word lists such as the B1 list for German (e.g., Glaboniat et al., 2016) focus largely on the spoken language of everyday conversations and interactions (Tschirner, 2019). Written language, however, is fundamentally different from the language of oral communication. The B1 list, for example, does not contain 40% of the 3,500 most frequent German words, according to the Routledge Frequency Dictionary of German (Jones & Tschirner, 2006; Tschirner et al., 2019). While the B1 list was derived from typical everyday encounters and linguistic demands, the corpus used as a basis for the frequency list was primarily in a written format.

In German, written-language text types display a greater word formation density, a higher average number of word formation components, a higher proportion of noun formations, and a higher proportion of explicit derivatives than spoken-language text types. While spoken-language text types exhibit a greater number of verb formations than written-language ones, this is due to the higher proportion of particle verbs rather than prefix verbs, the latter of which are more typical of writing (Stumpf, 2023). Niederhaus et al. (2016) examined seventh- and eighth-grade textbook chapters and tests focusing on fractions and discovered a prodigious amount of word formation products based on derivation and compounding. Tschirner (2019) found that almost 60% of the 8,000 most frequent words of German are based on word formation. Even listening tasks in school, such as listening to a teacher's explanations, require a great deal of academic vocabulary, which poses a significant challenge to monolingual and bilingual German elementary students alike. Heppt et al. (2014) found a large number of words consisting of three or more syllables, nominalizations, general academic words, and technical terms in listening passages for second- and third-grade students.

### 3. German Vocabulary Development in School Children

To the best of our knowledge, there has only been one study to date that has examined students' vocabulary skills in Romansh elementary schools in Graubünden. Gross & Imhof (2012) investigated the Romansh and German productive vocabulary range in written texts of fifth-grade students from monolingual Romansh and bilingual Romansh-German schools with the Guiraud (1954) index. They found that children with home languages other than Romansh or German, attending (oL1) ( $N = 40$ ) both Romansh and bilingual schools, had significantly lower Guiraud scores in their school language of Romansh or German respectively than children with L1 Romansh (RL1) and/or German (GL1) (both  $N = 216$ ) (Gross & Imhof 2012). The finding that oL1 students (often with L1 Portuguese) performed worse in the school language, let alone in German as their third and, in addition, linguistically distant language, is comparable to results from other Swiss studies. Analyzing the written texts of third- and fourth-grade students, Bonvin et al. (2018) determined that Portuguese-speaking children alphabetized in German or French schools showed lower productive lexical diversity<sup>1</sup> than their classmates with L1 German or French. Despite differences between the comparison groups and Portuguese-speaking students, competences overlapped considerably in both contexts. Furthermore, Bonvin et al. (2018) did not find any significant difference between Portuguese-speaking students and their comparison groups with regard to the corpus frequency of the words in the texts<sup>2</sup>. One reason for this may be that access to academic registers between the groups becomes more similar as the students progress through school.

<sup>1</sup> Lexical diversity refers to the proportion of words in a text that are not repetitions of words in the text (Bonvin et al., 2018).

<sup>2</sup> That is, the words they used have a similar frequency level, according to frequency corpora.

With regard to receptive vocabulary knowledge, similar levels of difference were found between oL1 students and GL1 students in German-speaking Switzerland (Röthlisberger et al., 2023), Austria (Seifert et al., 2019), and Germany (e.g., Lenhard & Lenhard, 2021). Röthlisberger et al. (2023) investigated longitudinal vocabulary breadth and depth in 322 GL1 and 51 oL1 Swiss students from second to third grade. Vocabulary breadth was measured auditorily with the German version of the standardized meaning-recognition PPVT-4 test (Lenhard et al., 2015). Vocabulary depth (defined as relational and semantic word knowledge by Juska-Bacher & Röthlisberger, 2021) was measured with antonym and hypernym recall items from the WWT-test (Wortschatz- und Wortfindungstest; Glück, 2021) and with meaning recall items from HAWIK-IV, the German version of the Wechsler Intelligence Test (Petermann & Petermann, 2007). GL1 students outperformed oL1 students in all results with large effect sizes (Röthlisberger et al., 2023).

Similarly, Seifert et al. (2019) found that GL1 students from first to third grade ( $N = 1,585$ ) in Austria were superior to oL1 students ( $N = 813$ ) in auditory receptive German vocabulary. Vocabulary size was measured with the 'Grazer Wortschatztest', a meaning recognition test with 30 word-picture matching items. In addition, oL1 students born in Austria performed significantly better than oL1 students born in other countries. Finally, Lenhard & Lenhard (2021) also found significant differences between GL1 and oL1 students in their norming of the German version of PPVT-4: Children and adolescents between 3 and 16 years of age living in Germany with a migration background (and thus often bilingual) performed significantly worse in the meaning recognition tests than monolingual German children. Children whose parents were both immigrants produced even lower results. While the difference between children with at least one immigrant parent and monolingual children decreased at school enrollment and leveled off in adolescence, children with two immigrant parents did not catch up until the age of 10 and remained constantly behind from that point on.

With respect to academic vocabulary skills in elementary school, there are only a few studies in the German context. Heppt et al. (2020) developed a specific written academic form recognition vocabulary test with versions for second and third/fourth grades. The 23 items of each version were selected from a criterion-based list of 118 academic words, which in turn was based on a frequency corpus of a German school context. In their validation study, they found that monolingual children from Germany ( $N = 1,667$ ) performed better than children with German and a second language as home languages ( $N = 1,122$ ) as well as children with home languages other than German ( $N = 836$ ).

Overall, the various studies with different vocabulary test instruments showed a difference between monolingual German and bilingual children, even when they had all been alphabetized in German-speaking schools. While there is one study of fifth-grade students in Graubünden (Gross & Imhof, 2012), the German academic vocabulary skills of students at the end of Romansh elementary education remains unclear.

#### 4. German in Elementary Schools in Graubünden

The present study examines the German vocabulary knowledge of students in Romansh schools in Switzerland. Romansh is a Romance language and one of the four official languages of Switzerland. As the historical language of the Swiss canton Graubünden, it is – along with Italian and German – one of the canton's official languages. Due to German migration, "the Romansh area and number of speakers of Romansh have been continually shrinking" (Gross, 2017, p. 10). As a result, only 21.5% (40,168) of the canton's population indicated Romansh as a 'regularly spoken' language in the last full census of 2000. Since Romansh is officially recognized as a minority language by the Charter for Regional or Minority Languages, the Swiss Confederation is committed to promote it in domains such as education, media and culture (Art. 8, 11 & 12 of the Europäische Charta der Regional- oder Minderheitensprachen of April 1st, 1998). Among these domains, compulsory education plays an especially important role in its preservation (Cathomas, 2005).

In Romansh schools in Graubünden, Romansh is the language of instruction in kindergarten and elementary school, while German is taught as a subject from third grade and English from fifth grade. In secondary school, the language of instruction is for the most part German (Deutschschweizer Erziehungsdirektoren-Konferenz, 2018). The purpose of this language model is to promote balanced, functional Romansh-German bilingualism by the end of compulsory school (Cathomas, 2005).

The Romansh school model appears to achieve its goals with respect to L1 Romansh and L1 German students: While L1 Romansh students perform significantly worse in general German tests than L1 German students at the end of Romansh elementary school, they perform similarly well in the same tests at the end of compulsory school (Cathomas, 2005). Nonetheless, we do not know much about students with other language

backgrounds. A considerable proportion of students in Romansh schools speak languages other than Romansh or German, predominantly Portuguese. In 2021, they accounted for 17% of all elementary school students (Bildungsstatistik Kanton Graubünden, 2022). According to Peyer et al. (2014), oL1 students perform as well in Romansh reading comprehension tests as their Romansh- and/or German-speaking classmates, but their results in German reading comprehension and written grammatical correctness are significantly worse. oL1 students from Romansh schools have much less contact with German outside of the school context than L1 Romansh students (Caglia et al., 2023; Prifti, 2016;). The apparent lack of German skills of oL1 students in Romansh schools seems to indicate a clear need for additional German language support. To classify their skills reliably, standardized tests are needed.

On account of German being the language of education from secondary school onwards, and the role vocabulary plays in reading and listening comprehension, we consider the assessment of students' general academic vocabulary to be crucial, particularly prior to the transition from elementary to secondary school. As demonstrated by Schuth et al. (2017), academic rather than general vocabulary significantly predicts children's academic performance. The instruments used in the studies cited in the previous chapter, however, appear to fall short for two reasons. First, none of them measure vocabulary ranges based on frequency. Although the academic vocabulary test of Heppt et al. (2020) was based on a frequency corpus, the final items were selected on a criterion-referenced basis and the number of items was too low to indicate a general vocabulary range. Secondly, none of the above tests were validated for the Swiss German context. The ITT Vocabulary Size Test (Institut für Testforschung und Testentwicklung, 2019), a frequency-based test developed on a large, predominantly written corpus, might be an alternative. This test has not yet been validated for elementary schools, let alone in a Romansh school setting. This study was designed therefore to explore the appropriateness of using the VST with fifth and sixth graders in Graubünden.

## 5. Research Questions

The following research questions guided our study:

- RQ1: How valid and reliable is the VST for 11- to 13-year-old elementary school children with German as L1, L2, or L3 in the context of Romansh schools in Switzerland?
- RQ2: How well does the VST distinguish between German as L1, Romansh as L1, and students with L1s other than German or Romansh?
- RQ3: What is the relationship between the VST and the Peabody Picture Vocabulary Test (PPVT)?

## 6. Methods

### 6.1 Participants

A total of 229 fifth- and sixth-grade students enrolled in 13 Romansh schools and one German elementary school in Graubünden, Switzerland, participated in the study. For reasons of anonymization, only classes from Romansh schools with at least two students who speak neither Romansh nor German at home were recruited. The German school with 35 students with L1 German served as a comparison group. For the present study, the data of three different test administrations with varying purposes was summarized<sup>3</sup>. The first test administration, in June 2021, was designed to pilot the VST with 44 Romansh elementary school students and to compare its results with the PPVT-4-Test for the main target group, i.e., oL1 students. The second test, from November 2021 to February 2022, was the pilot of another test in a study by Author 1 (Caglia, in prep.) and included 88 students from Romansh and German schools, while the third, from February to March 2022, was the pretest of an intervention study. It included 97 students from Romansh schools. 114 of the 229 participants were female and 115 were male. The mean age of the participants was 12.26 years with a minimum of 10.75, a maximum of 13.82, and a standard deviation of 0.68. The age of seven participants was unknown. Five participants were tested online individually under the supervision of the first author due to Covid restrictions. Table 1 shows the number of students who spoke German, Romansh, or other languages with their parent

<sup>3</sup> The various test times and test taker groups were combined to provide validity and reliability estimates of the VST independent of individual test administrations.

**Table 1***Student Language Background*

Language background	N
German (both parents)	76
Romansh (both parents)	52
Other (both parents)	43
German and Romansh	35
German and other	13
Romansh and other	10
<i>Total</i>	229

**6.2 Instruments and Procedure****6.2.1 Vocabulary Size Test (VST)**

The VST follows the model established by the English Vocabulary Levels Test (Nation, 2001), assessing how many of the 5,000 most frequent words of a language are known by the test taker. The German VST is based on the word frequency list of the Routledge Frequency Dictionary of German (Tschirner & Möhring, 2019). The test is divided into five bands of 1,000 words each: the most frequent 1,000; 1,001 to 2,000; 2,001 to 3,000; 3,001 to 4,000; and 4,001 to 5,000 words (Tschirner, 2021). For the present study, the receptive German VST 3 was used, which calculates the magnitude of a test taker's reading vocabulary.

Each band is sampled by 60 words, i.e., 30 nouns, 18 verbs, and 12 adjectives, randomly selected from the 1,000 words of the band to represent the whole band. They are allocated in ten part-of-speech clusters of six words each, for each of the five bands. Three of the words in each cluster are keys, i.e., they correspond to three definitions (synonyms, paraphrases, or gapped sentences), while the other three are additional distractors. For each definition, the same six multiple-choice options are provided, i.e., the six words of the cluster, from which the test taker selects an answer. For each band, the highest score is 30, and the total score for all five bands is 150 (Tschirner, 2021).

For the present study, the original German 3 VST was slightly modified for use in elementary school settings in Switzerland. For example, the letter 'ß' (Standard German from Germany) was replaced with 'ss' (Standard German from Switzerland), and the time limit was increased from 30 to 45 minutes. The pilot study in June 2021 revealed that the modified version of the German 3 VST performed as expected, and was used in all three test administrations.

Before taking the VST, the students filled out a questionnaire in the corresponding school language (German or Romansh). An explanation of the VST test procedure was then given in German or Romansh using sample slides (in 2022, an instructional video replaced the previous *ad hoc* explanations). Students entered their individual code online on the first page of the test. Once all codes had been checked by two people, i.e., the class teacher and Author 1 or a trained research assistant, a signal was given to start the test. When the students had completed the test, they were given an additional unrelated task to perform in silence.

After the June 2021 pilot, the test procedure was slightly adjusted in collaboration with Melanie Fuchs (University of Cologne). To make it more user-friendly for children, the ten clusters of each band were not presented on a single page that had to be scrolled but were shown on individual pages. The test progress was displayed on each page and a transitional page with a smiley was inserted after each band. In addition, students were given three practice items before the actual test. Finally, test results were not provided at the end. These practical optimizations had no statistical effect on the test results.<sup>4</sup>

<sup>4</sup> On average, students tested in June 2021 received only marginally lower scores ( $M = 100.98$ ,  $SE = 5.23$ ) than students tested in November 2021 and in 2022 ( $M = 109.02$ ,  $SE = 2.36$ ). A two-sided t-test revealed that this difference,  $-8.04$ , BCa 95% CI  $[-19.51, 3.43]$  was not significant,  $t(61.76) = -1.40$ ,  $p = 0.166$ . Equal variances were not assumed.



### 6.2.2 Peabody Picture Vocabulary Test

A subset ( $N = 17$ ) of the students who took the pilot test in June 2021 also took the German version of the Peabody Picture Vocabulary Test, 4th edition (Lenhard et al., 2015)<sup>5</sup>. All 17 students had at least one parent who did not speak Romansh or German natively, while both parents of 13 students spoke languages other than Romansh or German. This subsample, therefore, while small, provides an initial indication of the relationship between the PPVT and the VST of oL1 students.

The PPVT-4 measures the oral-receptive vocabulary knowledge of children and adolescents aged 3 to 17 years. It consists of 228 items in several different categories, such as animals or body parts. The items are evenly distributed over 19 sets (12 items each) in ascending difficulty (Lenhard & Lenhard, 2021). For each item, four color pictures are presented while the target word is pronounced by the test administrator. Younger children may point to the corresponding picture, while older children may name the corresponding number under the picture. The test administrator notes the answer on a recording sheet. The test is not timed and takes about 10-20 minutes.

The test procedure is adaptive and depends on the child's answers to the first set(s). After a warm-up and several practice items, the test administrator starts with a recommended start set, which depends on the child's age (e.g., start set 10 for children aged 12-13). If the child makes no more than one mistake, the start set becomes the floor set, and the test administrator continues to the next higher one. If the child makes two errors or more, the test administrator continues with one set below the start set and so on, until the floor set is determined. The test administrator then follows the numbered sequence. Once a child makes eight errors or more in a set, this set becomes the ceiling set (Lenhard et al., 2015).

Test results were normed by age group. The norming data were from a representative sample of 3,555 children and adolescents in Germany, ranging from 2 years and 7 months up to 18 years. 29.3% of the sample had a migration background. Test reliability was very high: EAP reliability  $r = .965$ ; odd-even-split-half reliability at least  $r = .92$ ; retest reliability  $r = .91$  (Lenhard & Lenhard, 2021).

While the reliability of the German version of the PPVT-4 was shown to be high, convergent validity measures were lower. For sixth- and seventh-grade students ( $N = 67$ ), the correlation between the PPVT-4 and WISC-IV test assessing productive meaning recall, which includes abstract words, was moderate:  $r = 0.66$ , while the correlation with ELFE 1-6, a reading comprehension test administered to second- to fourth-grade students ( $N = 46$ ), was relatively low:  $r = 0.44$  (Lenhard et al., 2015).

In the present study, the first author piloted an adapted PPVT-4 procedure of ten sixth-grade children with different home languages in a Romansh school in May 2021, and then used this procedure in June 2021. 13 of the 17 students spoke neither Romansh nor German at home, two spoke Romansh and another two spoke German with one parent at home. After the administration of the VST and a 20-minute break, the PPVT-4-test was administered individually by the first author and a trained research assistant. Romansh was the language of communication during testing. In the test, facilitators read the German target words according to the standard PPVT-4 testing protocol. Students were allowed to call out the number of the matching picture in either Romansh or German. Because some of the children tested (aged 12-13) might not have much contact with German outside of school (see section 4), different start sets were used, depending on the students' linguistic background (see Table 2).

**Table 2**

*Start Sets Used in the Present Study*

Language spoken with parents	Start set used in the present study	Recommended age of start set (Lenhard et al. 2015)
Ger-Ger* / Ger-Rm	10	12-13 years
Rm-Rm	9	10-11 years
Rm-o / Ger-o	7	8 years
o-o	5	6 years

\*Note: Ger = German; Rm = Romansh; o = other languages

<sup>5</sup> Because the VST performed successfully with the target group in its first use in June 2021, the much more time-consuming PPVT-4 test was not used in subsequent testing.

Unlike Juska-Bacher & Röthlisberger (2021), German words that are typical in Germany but not in Switzerland (e.g., ‘Schornstein’ vs. ‘Kamin’; ‘Umschlag’ vs. ‘Kuvert’; ‘Möhre’ vs. ‘Karotte’/‘Rüebli’) were not changed. The reason for this was that the test was not adapted and standardized for the Swiss context at that time, and we wanted to maximize comparability with the studies in Germany.

## 7. Results

To provide evidence of the reliability and validity of the VST in Romansh elementary school settings (RQ1), Cronbach’s alpha was computed for the complete test and for each of the five bands. In addition, group means were examined to determine whether there were any differences between fifth-grade and sixth-grade students, and between students who spoke German, Romansh, and/or other languages at home. Finally, the correlation between the VST and the PPVT-4 was calculated to determine the relationship between the tests. While Cronbach’s alpha provides evidence of the internal consistency of the test, which may be interpreted as a measure of reliability and internal validity, the group mean differences and correlation results may be interpreted as evidence of concurrent validity, i.e., independent measures supporting the argument that the VST measures actual differences in the vocabulary size of students.

### 7.1 Internal Consistency

The complete test consists of 150 items, 30 in each of the five bands. Correct responses were coded as 1 and incorrect responses as 0. Items that were not attempted were left blank. Table 3 documents the descriptive statistics of the total score.

**Table 3**

*Descriptive Statistics of the Total Score of the German 3 VST*

N	Min	Max	Mean	S.E.	SD	Skewness	S.E.	Kurtosis	S.E.
229	5	148	107.48	2.16	32.72	-0.96	0.16	-0.10	0.32

Table 3 shows that there were 229 participants with total scores ranging from 5 to 148 out of 150, all but covering the complete spread of scores. The mean and standard deviation demonstrate that most students had relatively high vocabulary sizes. Skewness and kurtosis values between -2 and 2 suggest a reasonably normal distribution (George & Mallery, 2010).

Total scores are more fine-grained than band scores based on 80% correct answers because there are no hard borders. Total scores were therefore used instead of bands to evaluate the validity and reliability of the receptive German 3 VST. To examine the overall reliability, Cronbach’s alpha was calculated between the five bands of the test. Cronbach’s alpha measures the similarity of the results of the different bands. If alpha is high, it may be assumed that all items measure the same construct, in this case receptive vocabulary size. Cronbach’s alpha above 0.7 is considered acceptable, above 0.8, it is considered good, and above 0.9 very good. Table 4 shows the number of tests administered, Cronbach’s alpha, and the number of items, in this case bands.

**Table 4**

*Cronbach’s Alpha of the Receptive German 3 VST*

N of Tests	Cronbach’s Alpha	N of Items
229	0.96	5

Table 4 shows that the reliability and internal validity of the receptive German 3 VST was above 0.9, which supports the claim that the test is valid and reliable. Table 5 displays the item statistics of the five bands. Note that *N* is different from the total number of tests because the calculations are based on students who answered every item in a band, i.e., did not leave any items blank.

**Table 5***Means and Standard Deviations of the Five Bands of the Receptive German 3 VST*

Band	Mean	SD	N
1000	24.62	4.87	191
2000	23.67	5.12	193
3000	22.73	7.29	179
4000	21.29	7.08	191
5000	20.55	7.28	179

Table 5 demonstrates that the mean decreases as the bands go up, which indicates that bands are increasing in difficulty. This also provides evidence that the test measures vocabulary size, i.e., test takers know more of the more frequent words than of the less frequent ones. To examine the internal consistency of each band, Cronbach's alpha was calculated for each band of 1,000 words. Table 6 displays the number of test takers, Cronbach's alpha, and the number of items for each band.

**Table 6***Cronbach's Alpha for Each Band of the Receptive German 3 VST*

Band	N of Tests	Alpha	N of Items
1	191	0.88	30
2	193	0.88	30
3	179	0.93	30
4	191	0.92	30
5	179	0.92	30
Total	229		

Table 6 demonstrates that the internal consistency of each band was close to or above 0.9, i.e., close to or within the very good range for all five bands.

Cronbach's alpha between the five bands of the test and between the 30 items of each band, together with the fact that the mean decreases as the words become less frequent, provide strong evidence of the internal consistency of the German 3 VST in a Swiss elementary school setting.

## 7.2 Student Standing and Language Background

Two sets of group means were examined to provide evidence of concurrent validity by investigating the relationship between the VST results and two external variables: student standing and language background. Students in sixth grade are expected to have a bigger vocabulary size than students in fifth grade, while students who speak German with both parents should have a bigger vocabulary size than, for example, students who speak a language other than German at home.<sup>6</sup> Table 7 shows the number of students in fifth and sixth grade, their mean total score, standard deviation, and standard error of the mean.

**Table 7***Total Vocabulary Score by Grade*

Grade	N	Mean	SD	S.E.
6	137	112.43	31.28	2.67

<sup>6</sup> The variable *school type*, which is not critical to the present argument, is considered in another article (Caglia & Tschirner, in prep.).



5	92	100.1	33.58	3.50
---	----	-------	-------	------

Table 7 documents that 137 students were in sixth grade, while 92 were in fifth grade. The mean total vocabulary score of sixth-grade students was 112, and 100 for fifth-grade students. An independent samples t-test revealed that the difference in means was significant,  $t(227) = 2.84$ ,  $p = 0.005$ , indicating that the vocabulary size of the sixth-grade students was, indeed, larger than that of the fifth-grade students. Cohen's  $d$  was 0.38, indicating a small to medium effect size.

To examine language background, students were grouped according to their home language(s), i.e., the language(s) they spoke with their parents most of the time. Students were divided into six groups. Table 8 documents the number of students in each group, the mean total score of the group, and the standard deviation.

**Table 8**

*Home Language Groups*

Home Language	N	Mean	SD
Other-Other	43	55.98	23.97
Other-Romansh	10	87.70	22.41
Other-German	13	125.31	17.40
Romansh-Romansh	52	115.12	20.80
Romansh-German	35	118.83	18.01
German-German	76	125.71	17.81
<i>Total</i>	<i>229</i>	<i>107.48</i>	<i>32.72</i>

Note 1-6: Languages spoken with parents most of the time. Students who spoke two languages with one parent equally often were coded as follows: German if one of the languages spoken was German; Romansh if one of the languages spoken was Romansh, while the other one was a language other than German or Romansh.

Table 8 shows that the groups with the lowest means were O-O (home language other than Romansh or German) and O-R (one home language Romansh and one other than Romansh or German), while the other four groups had similar means. An ANOVA was run to examine group means more fully. Levene's test of equality of error variances was non-significant ( $p = 0.58$ ), indicating that the error variance of the dependent variable was equal across groups. The ANOVA demonstrated that there was a significant and large effect of group membership on total score,  $F(5, 223) = 77.49$ ,  $p < 0.001$ ,  $\eta^2 = 0.63$ . Table 9 shows the results of the post-hoc Scheffé test.

**Table 9**

*Post-hoc Scheffé Test of the Effect of Group Membership on Total Score*

Group	N	1	2	3
O-O	43	55.98		
O-R	10		87.70	
R-R	52			115.12
R-G	35			118.83
O-G	13			125.31
G-G	76			125.71

Note: Alpha = 0.05; Home Languages: G = German, R = Romansh, O = Language other than Romansh or German.

Table 9 demonstrates that there were three significantly different groups: Group O-O (home language(s) other than German or Romansh), Group O-R (one home language is different from German or Romansh, while the other is Romansh), and all other groups. The effect size was large ( $\eta^2 = 0.63$ ), indicating that language background had a significant influence on the size of student vocabulary. In addition, the differences in means between the three groups were considerable. Group O-O had a mean total score of 56, roughly equivalent to

a vocabulary size of 1,000 words, while Group O-R had a mean total score of 88, equivalent to a vocabulary size of 3,000 words. Group O-G had mean scores between 115 and 126, equivalent to 4,000 and 5,000 words respectively.

There were relatively few students in Groups O-R ( $N = 10$ ) and O-G ( $N = 13$ ), while Groups O-O, R-R, R-G and G-G had between 35 and 77 students. Focusing on the larger groups, the results indicate that students whose home language was other than German or Romansh had a relatively low vocabulary size (1,000), while students who spoke German or Romansh at home had 4,000 words or more. It is interesting to note that there were no significant differences between students with Romansh/Romansh, German/Romansh, and German/German home languages and that all three groups scored near the top of the scale (between 115 and 126).

These results support the assumption that the German 3 VST measures actual vocabulary size because it distinguishes significantly between students whose home language is other than German or Romansh and students who speak German or Romansh at home, with the latter having considerably larger vocabulary sizes.

#### Relationship Between the VST and the PPVT-4

To provide further evidence of external validity, the relationship between the German VST 3 and the PPVT-4 was investigated. While the PPVT-4 is a translation of the English version with similar word frequencies and difficulties and is limited to words that may be visualized (Lenhard & Lenhard, 2021), the VST, by its design, is based on a German frequency corpus, which commonly includes about 75% of written texts and 25% oral passages. Nevertheless, if the VST measures vocabulary size adequately, there should be a relationship between the two tests. Table 10 shows the descriptive statistics of the subgroup of students from a Romansh school who took both the VST and the PPVT-4. All 17 students were in sixth grade; 13 of them spoke a language other than German or Romansh with both parents, while four spoke German or Romansh with one parent and a different language with the other parent.

**Table 10**

*Descriptive Statistics of the VST Total Score and the Results of the PPVT-4*

Test	N	Min	Max	Median	Mean	S.E.	SD
VST	17	5	135	73.00	69.47	6.67	27.49
PPVT-4	17	77	187	126	131.00	6.12	25.35

Table 10 documents that both mean VST and PPVT scores were approximately in the middle of the scale (maximum VST score: 150; maximum PPVT score: 228). Table 11 shows the correlation between the tests.

**Table 11**

*Correlation Between the VST and the PPVT-4*

N	Pearson's r	Spearman's rho
17	0.81**	0.74**

\*\*  $p < 0.01$

Table 11 documents that the correlation between the tests was significant and strong,  $r = 0.81$ ,  $\rho = 0.74$ ,  $p < 0.01$ . Both tests explained approx. 66% of each other's results ( $r^2 = 0.66$ ). This provides further evidence of the claim that the VST measures actual vocabulary size. It also indicates that about 34% of the relationship remains unexplained. This is most likely due to the different approach to the vocabulary corpus with the VST being a purely frequency-based test and the PPVT-4 focusing on a translation that is as close to English as possible and limited to a visualizable word corpus.

## 8. Conclusion

Before summarizing the results of our study, we would like to point out some limitations. First, the subsample tested with the PPVT-4-Test was small ( $N = 17$ ) and not represented by home and school language. In addition, the PPVT-4 test was standardized in Germany and might underestimate vocabulary knowledge in the Swiss context. As for the VST test, its procedure was slightly adapted between the first and subsequent tests. However, t-tests showed that this adaption had only a minimal influence on test results (see footnote 4).

The present study showed that the receptive German 3 VST is very reliable for fifth- and sixth-grade Romansh elementary school students in Graubünden, Switzerland. It shows good to very good internal consistency, as Cronbach's alpha is between 0.88 and 0.93 and the mean values of the frequency bands decrease with increasing difficulty. Although the subsample tested in the PPVT-4 was small ( $N = 17$ ), the VST correlated strongly ( $r = 0.81$ ) and significantly with the PPVT-4. The strong correlation with a test developed and standardized for children suggests that the VST is appropriate for elementary school children in grades five and six. Furthermore, the VST was able to distinguish between several groups: children who spoke neither Romansh nor German at home, children who spoke Romansh and another language at home, and all other groups. Children who spoke another language with both parents had a relatively low vocabulary. This result is in line with previous studies (Gross & Imhof, 2012; Peyer et al., 2014). As the test has performed well with L1 German students in a German school, it might also be useful in other German schools in Switzerland<sup>7</sup>. However, representative samples of additional students who speak languages other than German and of students in other cantons in Switzerland would be valuable.

In terms of ecological validity, the VST may be considered an internationally valid instrument, as the test is based on a representative German corpus from Germany, Austria, and Switzerland. Because it is based on an adult and predominantly written corpus, it is measuring academic rather than general vocabulary skills. This focus is informative in contexts where German is or about to be the language of education. However, as there is no frequency corpus to date focusing on Swiss secondary schools, it remains unclear how well the VST actually covers academic vocabulary needs in this specific context. The *BiSpra* German academic written and oral vocabulary corpus focuses on the language used in elementary schools in Germany. While there is no frequency dictionary based on that corpus, Redder & Weinert (2015) published a criterion-based wordlist with general academic words selected from it. A comparison of the wordlist with the frequency dictionary on which the VST is based (Tschirner & Möhring, 2019) revealed that 110 of the 118 words from Redder & Weinert's (2015) list (93%) were among the 5,000 most frequent words from the adult corpus. This suggests that the VST may cover a considerable portion of general academic vocabulary, useful in elementary school contexts.

## Bibliography

- Bildungsstatistik Kanton Graubünden. (2022). *Anzahl Schülerinnen und Schüler der öffentlichen Volksschule, Kanton Graubünden, Schuljahr 2021/22. Nach Schulstandort, Erstsprache und Stufe* [anonymisierte Excel-Datentabelle]. <https://phgr.sharepoint.com/:x/s/public/fed1/EbdC-eoBJhJOil8qPrni2eoBtp6N3uWeMZfcHOhr3K4PA?e=INMUO5>
- Bonvin, A., Vanhove, J., Berthele, R., & Lambelet, A. (2018). Die Entwicklung von produktiven lexikalischen Kompetenzen bei Schüler(innen) mit portugiesischem Migrationshintergrund in der Schweiz. *Zeitschrift Für Interkulturellen Fremdsprachenunterricht. Didaktik Und Methodik Im Bereich Deutsch Als Fremdsprache*, 23(1), 135–148. <https://zif.tu-journals.ulb.tu-darmstadt.de/article/id/3146/>
- Caglia, D. (in prep.). *Eine Minderheit in der Minderheit – Wortschatzförderung im Deutschen für Schüler:innen mit heterogenen Deutschkenntnissen an rätoromanischen Schulen* [dissertation project].
- Caglia, D., & Tschirner, E. (in prep.). *German vocabulary sizes of bilingual and monolingual elementary school students in Graubünden, Switzerland*.
- Caglia, D., Graf, F., & Ke, K. (2023). *Bildungschancen durch Mehrsprachigkeit an romanischsprachigen Volksschulen: Bericht Untersuchungsgemeinde 1. Situationsanalyse und Massnahmenvorschläge*. Pädagogische Hochschule Graubünden; Institut für Kulturforschung Graubünden.
- Cathomas, R. (2005). *Schule und Zweisprachigkeit: Immersiver Unterricht: internationaler Forschungsstand und eine empirische Untersuchung am Beispiel des rätoromanisch-deutschen Schulmodells in der Schweiz*. Waxmann.
- Europäische Charta der Regional- oder Minderheitensprachen of April 1st, 1998 (AS 2003 2507; Status on September 26th, 2024).
- George, D., & Mallery, P. (2010). *Spss for Windows step by step: A simple guide and reference, 17.0 update* (10<sup>th</sup> ed.). Boston.
- Glaboniat, M., Perlmann-Balme, M., & Studer, T. (2016). *Goethe-Zertifikat B1. Deutschprüfung für Jugendliche und Erwachsene. Wortliste*. Goethe-Institut.

<sup>7</sup> The VST has also been used recently with secondary school students in Germany (Marx et al., 2024).

- Glück, C. W. (2021). Der WWT 6-10 und seine Testitems. *Bulletin Suisse de Linguistique Appliquée*, 113, 29-47. <https://www.e-periodica.ch/digbib/view?pid=val-003%3A2021%3A0%3A%3A4#37>
- Gross, M. (2017). *Romansh: The Romansh language in education in Switzerland*. Regional Dossiers Series. Mercator European Research Centre on Multilingualism and Language Learning. [https://www.mercator-research.eu/fileadmin/mercator/documents/regional\\_dossiers/romansh\\_in\\_switzerland.pdf](https://www.mercator-research.eu/fileadmin/mercator/documents/regional_dossiers/romansh_in_switzerland.pdf)
- Gross, M., & Imhof, A. (2012). Entwicklungsstand des Wortschatzes. In M. Gross, & L. Flepp (Eds.), *Collana PHGR: Vol. 1. Mamma eu les cuelas chochas: Schreibkompetenzstudie Graubünden 5. Primarschulklasse = una ricerca sulle competenze di scrittura nel Grigioni romanciofono e italofono* (pp. 97-105). Südostschweiz-Buchverlag.
- Guiraud, P. (1954). *Les caractères statistiques du vocabulaire : Essai de méthodologie*. Presses universitaires de France.
- Hacking, J. F., Rubio, F., & Tschirner, E. (2019). Vocabulary Size, Reading Proficiency and Curricular Design: The Case of College Chinese, Russian and Spanish. In P. Winke, & S. M. Gass (Eds.), *Educational Linguistics. Foreign Language Proficiency in Higher Education* (Vol. 37, pp. 25-44). Springer International Publishing. [https://doi.org/10.1007/978-3-030-01006-5\\_3](https://doi.org/10.1007/978-3-030-01006-5_3)
- Heppt, B., Köhne-Fuetterer, J., Eglinsky, J., Volodina, A., Stanat, P., & Weinert, S. (2020). *BiSpra 2-4: Test zur Erfassung bildungssprachlicher Kompetenzen bei Grundschulkindern der Jahrgangsstufen 2 bis 4* (1<sup>st</sup> ed.). Bildungsbezogene Sprachtests: Vol. 2. Waxmann.
- Heppt, B., Stanat, P., Dragon, N., Berendes, K., & Weinert, S. (2014). Bildungssprachliche Anforderungen und Hörverstehen bei Kindern mit deutscher und nicht-deutscher Familiensprache. *Zeitschrift Für Pädagogische Psychologie*, 28(3), 139-149. <https://doi.org/10.1024/1010-0652/a000130>
- Huhta, A., Alderson, J. C., Nieminen, L., & Ullakonoja, R. (2011, August 4). *Diagnosing reading in L2: Predictors and vocabulary profiles* [Conference presentation]. ACTFL CEFR Conference, Provo UT, United States.
- Institut für Testforschung und Testentwicklung. (2019). *Wortschatztests*. Institut für Testforschung und Testentwicklung. <https://itt-leipzig.de/wortschatztests/>
- Jones, R. L., & Tschirner, E. (2006). *A frequency dictionary of German: Core vocabulary for learners*. Routledge frequency dictionaries. Routledge. <http://www.loc.gov/catdir/enhancements/fy0653/2005012949-d.html>
- Juska-Bacher, B., & Röthlisberger, M. (2021). Das Konstrukt Wortschatz und seine Dimension(en) Umfang und Tiefe? Empirische Ergebnisse aus der Unterstufe. *Bulletin Suisse de Linguistique Appliquée*, 113, 49-68. <https://www.e-periodica.ch/digbib/view?pid=val-003%3A2021%3A0%3A%3A4#57>
- Deutschschweizer Erziehungsdirektoren-Konferenz. (2018). *Lektionentafeln Volksschule GR ab Schuljahr 2018/19*. Lehrplan 21 GR. [https://lehrplan21.ch/sites/default/files/GR\\_Lehrplan21\\_Lektionentafeln\\_ab\\_2018\\_19.pdf](https://lehrplan21.ch/sites/default/files/GR_Lehrplan21_Lektionentafeln_ab_2018_19.pdf)
- Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test - Revision 4 (PPVT-4), deutsche Version*. Pearson Assessment.
- Lenhard, W., & Lenhard, A. (2021). Bedeutung und Diagnostik des Wortschatzes am Beispiel des Peabody Picture Vocabulary Test (PPVT-IV). *Bulletin Suisse de Linguistique Appliquée*, 113, 9-27. <https://www.e-periodica.ch/digbib/view?pid=val-003%3A2021%3A0%3A%3A4#17>
- Marx, N., Barberio, T., Twente, L. R., Fuchs, M., Eisenbeiss, S., Dewitz, N. von, & Bredthauer, S. (2024). *Abschlussbericht zum Projekt „Sprachkompetenzen neuzugewandelter Schülerinnen und Schüler im Regelunterricht“ (01.2021 - 06.2024)* [Research report]. University of Cologne. <http://kups.ub.uni-koeln.de/id/eprint/73168>
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters. <https://doi.org/10.21832/9781847692092>
- Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Bartning, & M. Martin (Eds.), *EUROSLA monographs series: Vol. 1. Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 211-232). European Second Language Association.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Niederhaus, C., Pöhler, B., & Prediger, S. (2016). Relevante Sprachmittel für mathematische Textaufgaben – Korpuslinguistische Annäherung am Beispiel Prozentrechnung. In E. Tschirner, Bärenfänger, O., & Möhring, J. (Eds.), *Kompetenzprofile Deutsch als fremde Bildungssprache: Das Spannungsfeld von Fachwissen, sprachlicher Kompetenz, Diagnostik und Didaktik* (pp. 135-162). Stauffenburg.
- Petermann, F., & Petermann, U. (2007). *HAWIK-IV Hamburg-Wechsler-Intelligenztest für Kinder – IV. Übersetzung und Adaptation des WISC-IV von D. Wechsler*. Huber.
- Peyer, E., Lindt-Bangerter, B., Graber, S., & Camenisch, S. (2014). *Projektbericht des SNF-Projekts „Empfehlungen für Basisstandards für die Schulsprachen der Rätromanen“*. Universität Freiburg, Schweiz. <https://folia.unifr.ch/unifr/documents/303833>
- Prifti, E. (2016). Dynamiken des Sprachverhaltens bei portugiesischen Migranten im Engadin: Eine pluridimensionale migration-linguistische Analyse. In J. Born, & A. Ladilova (Eds.), *Iberolinguistica: Vol. 2. Sprachkontakte des Portugiesischen: Ausgewählte Beiträge der gleichnamigen Sektion des 10. Deutschen Lusitanistentages vom 10.-14. September 2013 in Hamburg* (pp. 247-275). Peter Lang GmbH Internationaler Verlag der Wissenschaften.
- Redder, A., & Weinert, S. (2015). Hamburg-Bamberger BiSpra-Liste. In A. Redder, J. Naumann, R. Tracy, & S. Lambert (Eds.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung: Ergebnisse* (pp. 90-92). Waxmann.
- Röthlisberger, M., Zanger, C., & Juska-Bacher, B. (2023). The role of vocabulary components in second language learners' early reading comprehension. *Journal of Research in Reading*, 46(1), 1-21. <https://doi.org/10.1111/1467-9817.12411>
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363. <https://doi.org/10.1177/1362168808089921>
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The Percentage of Words Known in a Text and Reading Comprehension. *The Modern Language Journal*, 95(1), 26-43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction*, 49, 157-165. <https://doi.org/10.1016/j.learninstruc.2017.01.005>

- Seifert, S., Paleczek, L., & Gasteiger-Klicpera, B. (2019). Rezeptive Wortschatzleistungen in der Grundschule. Unterschiede zwischen Kindern mit Deutsch als Erst- oder Zweitsprache. *Empirische Sonderpädagogik*, 11(4), 259–278. <https://doi.org/10.25656/01:18334>
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal*, 36(2), 139–152. <https://doi.org/10.1080/09571730802389975>
- Stumpf, S. (2023). *Wortbildung diamedial: Korpusstudien zum geschriebenen und gesprochenen Deutsch*. De Gruyter.
- Tschirner, E. (2021). *Examining the Validity and Reliability of the ITT Vocabulary Size Tests* (Research Papers in Assessment 3). Institut für Testforschung und Testentwicklung. <http://nbn-resolving.de/urn:nbn:de:bsz:15-qucosa2-760957>
- Tschirner, E. (2019). Der rezeptive Wortschatzbedarf im Deutschen als Fremdsprache. In E. Peyer, T. Studer, & I. Thonhauser (Eds.), *IDT 2017. Band 1: Hauptvorträge* (pp. 98–111). Erich Schmidt.
- Tschirner, E. (2009). Korpuslinguistik und Fremdsprachenunterricht. In I. Bartoszewicz, M. Hałub, & E. Tomiczek (Eds.), *Argumente-Profil-Synthesen* [Germanica Wratislaviensia 129] (pp. 117–132). Wydawnictwo Uniwersytetu Wrocławskiego.
- Tschirner, E., Hacking, J., & Rubio, F. (2019). Reading proficiency and vocabulary size: An empirical investigation. In P. Ecke, & S. Rott (Eds.), *Understanding vocabulary learning and teaching: Implications for language program development* (pp. 58–77). Cengage.
- Tschirner, E., & Möhring, J. (2019). *Frequency dictionary of German: Core vocabulary for learners* (2<sup>nd</sup> ed.). Routledge.

**Keywords:** German as a foreign language; German as a second language; elementary school; academic vocabulary; test validity

## Schätzung des deutschen Wortschatzumfangs von Primarschüler\*innen in Graubünden, Schweiz

### Zusammenfassung

Wortschatz korreliert stark mit generellen Sprachfähigkeiten und ist ein guter Prädiktor für bildungsrelevante Kompetenzen. Besonders aufschlussreich sind frequenzbasierte Wortschatztests, wie etwa der deutsche VST ('Vocabulary Size Test'; Institut für Testforschung und Testentwicklung, 2019; Nation, 2001). In der vorliegenden Studie wurde dieser Test mit 229 Fünft- und Sechstklässler\*innen aus rätoromanischen und deutschsprachigen Primarschulen in Graubünden (Schweiz) validiert. Die Ergebnisse zeigen eine hohe Reliabilität und hohe interne Validität des deutschen VST sowie eine ausgeprägte Korrelation des Tests mit dem PPVT-4 (Lenhard et al., 2015) bei einer Teilstichprobe. Damit bietet der Test relevante Daten zum bildungssprachlichen Wortschatz von Schüler\*innen mit Deutsch als L1, L2 oder L3 im rätoromanischen Kontext.

**Schlagworte:** Deutsch als Fremdsprache; Deutsch als Zweitsprache; Primarschule; bildungssprachlicher Wortschatz; Testvalidität

## Estimation du vocabulaire allemand des élèves du primaire dans les Grisons, Suisse

### Résumé

Le vocabulaire est fortement corrélé aux compétences linguistiques générales et permet de prédire les compétences académiques. Les tests de vocabulaire basés sur la fréquence, comme le VST allemand ('Vocabulary Size Test'; Institut für Testforschung und Testentwicklung, 2019; Nation, 2001), sont particulièrement informatifs. Cette étude a validé ce test avec 229 élèves de cinquième et sixième année d'écoles primaires romanches et allemandes des Grisons, en Suisse. L'étude montre une fiabilité et une validité interne élevées du VST allemand et une forte corrélation du test avec le PPVT-4 (Lenhard et al., 2015) dans un sous-échantillon. Le test offre ainsi des données pertinentes sur le vocabulaire académique des élèves ayant l'allemand comme L1, L2 ou L3 dans le contexte romanche.

**Mots-clefs :** allemand comme langue étrangère ; allemand comme langue seconde ; école primaire ; vocabulaire académique ; validité des tests



## Stima del vocabolario tedesco delle allieve/degli allievi della scuola primaria dei Grigioni, Svizzera

### Riassunto

Il vocabolario è fortemente correlato alle competenze linguistiche generali ed è predittivo delle competenze accademiche. I test di vocabolario basati sulla frequenza, come il VST tedesco ('Vocabulary Size Test'; Institut für Testforschung und Testentwicklung, 2019; Nation, 2001), sono particolarmente informativi. Il presente studio ha validato questo test con 229 allievi di quinta e sesta classe delle scuole primarie romance e tedesche dei Grigioni, in Svizzera. Lo studio mostra un'elevata affidabilità e validità interna del VST tedesco, nonché una forte correlazione del test con il PPVT-4 (Lenhard et al., 2015) in un sottocampione. Il test offre pertanto dati rilevanti sul vocabolario accademico di allieve e allievi con il tedesco come L1, L2 o L3 nel contesto romancio.

**Parole chiave:** tedesco come lingua straniera; tedesco come lingua seconda; scuola primaria; vocabolario accademico; validità dei test

### Resumaziun

La grondezza dal vocabulari correlescha fermamain cun las cumpetenzas linguisticas generalas e predi las cumpetenzas academicas. Ils tests da vocabulari che sa basan sin la frequenza, sco il VST tudestg ('Vocabulary Size Test'; Institut für Testforschung und Testentwicklung, 2019; Nation, 2001), èn particularmain infurmativs. Il studi preschent ha validà quest test cun 229 scolar\*as da las tschintg- e sisavlas classas da scolas primaras rumantschas e tudestgas dal Grischun en Svizra. Il studi mussa in'auto reliabilitad e validitad interna dal VST tudestg ed ina ferma correlaziun dal test cun il PPVT-4 (Lenhard et al., 2015) en ina sutgruppa. Il test porscha pia datas relevantas davart il vocabulari academic da scolar\*as cun il tudestg sco L1, L2 u L3 en il context rumantsch.

**Pleds-clav:** Tudestg sco lingua estra; tudestg sco segunda lingua; scola primara; vocabulari academic; validitad da tests

**Dominique Caglia** (n. Dosch), doctoral student in German as a foreign and second language at the University of Fribourg. Studied pedagogy/psychology and Romansh and conducts research at the PH Graubünden on multilingualism in schools.

Pädagogische Hochschule Graubünden, Scalärastrasse 17, CH-7000 Chur

E-Mail: [dominique.caglia@phgr.ch](mailto:dominique.caglia@phgr.ch)

**Erwin Tschirner**, Gerhard-Helbig-Professor of German as a Foreign Language, Emeritus at the University of Leipzig, Germany; main research areas: Second Language Acquisition (speaking, reading, listening, vocabulary), language proficiency assessment and assessment theory, German syntax and morphology, language teaching.

Universität Leipzig, Herder Institut, Beethovenstrasse 15, D-04107 Leipzig

E-Mail: [tschirner@uni-leipzig.de](mailto:tschirner@uni-leipzig.de)