
**ITT Technical Reports on Language Testing 3**

# Assigning Frequency Bands to the Productive Vocabulary Size Test According to the Total Score of the Test Taker

Erwin Tschirner[1]

[1]Leipzig University, Germany

**Author Note**

Erwin Tschirner https://orcid.org/0000-0002-5915-5344

Correspondence concerning this report should be addressed to Erwin Tschirner, Herder-Institut, Universität Leipzig, Beethovenstraße 15, 04107 Leipzig. email: tschirner@uni-leipzig.de

Bibliographische Informationen der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Die "ITT Technical Reports on Language Testing" sind eine Reihe des Instituts für Testforschung und Testentwicklung e. V. (ITT), in der Technische Reports zu unterschiedlichen Sprachtestevaluationen veröffentlicht werden.

Institut für Testforschung und Testentwicklung e.V. Leipzig
c/o Herder-Institut
Universität Leipzig
Beethovenstraße 15
04107 Leipzig
www.itt-leipzig.de

Herausgeberschaft:
Olaf Bärenfänger, Universität Leipzig
Jupp Möhring, Technische Universität Dresden
Erwin Tschirner, Universität Leipzig

Redaktion:
Lisa Lenort, Elisabeth Muntschick, Antonia Kurz




URN des Bandes: https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-925275
URN der Reihe: https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-902765

ISSN **2942-741X**

**Table of Contents**



**List of Tables**

**List of Figures**

## Methods and Analysis

The Productive Vocabulary Size Test (P-VST) measures how many of the most frequent 5,000 words of a language are controlled productively by a test taker. It consists of five bands of 1,000 words each, involving 18 gapped sentences with the target word missing except for one or several initial letters to ensure only one word fits the sentence context accurately. It is commonly scored according to how many of the words of each band are used correctly when writing, usually around 70%-80% correct. Tschirner and Möhring (2024) argued – based on a very large dataset – that 70% may be the more appropriate percentage. This is the percentage used in the present study.

The highest of the five bands that has at least 70% correct is considered the vocabulary size of the test taker. Each band has a maximum score of 18; 70% correct, therefore, amounts to a score of 13. These cut scores, however, increase the possibility that the result may be unduly influenced by chance. Imagine, e.g., a test taker who has a score of 12 on all five bands, whose final rating would, therefore, be "less than 1000 words." Moreover, reducing the results of each band to a yes/no decision, does not utilize the richness of the original scale, i.e., 90 individual items. Furthermore, an estimation of someone's vocabulary size based on total score may be done more rapidly than having to score each band separately and combining the results. The aim of this analysis, therefore, was to use the total score rather than the 70% correct approach to predict vocabulary size.

To examine the relationship between total scores and vocabulary size, all productive VSTs administered to a total of 1,241 test takers who had been part of vocabulary studies under controlled conditions were analyzed. These included 623 test takers at a university in Germany (see, e.g., Möhring 2022, Wei et al.,2023) and 618 examinees at a US university (see, e.g., Spino-Seijas & Tschirner, 2023). Table 1 shows the languages involved.

**Table 1**

*Languages Involved in the Analysis*

| | Frequency | Percent |
|---|---|---|
| Chinese | 41 | 3.3 |
| French | 88 | 7.1 |
| German | 671 | 54.1 |
| Italian | 95 | 7.7 |
| Japanese | 52 | 4.2 |
| Spanish | 294 | 23.7 |
| Total | 1241 | 100 |

The great majority of the German test takers studied in Germany and had a variety of language backgrounds. All other test takers were either monolingual or bilingual English speakers. For many of these test takers, reading, listening, and speaking proficiency data were also available and were used to provide evidence of convergent validity.

To establish total score cut scores for bands, descriptive statistics for all test takers who were assigned a particular band according to the 70% correct criterion were calculated.[1] Table 2 shows the bands; the number of test takers who attained the band; the minimum and maximum, median and mean total score; the standard error of the mean; the standard deviation; and the mean minus 1.5 times as well as 2 times the standard deviation. The mean of the test takers at each band minus 1.5

[1] While languages may differ in the vocabulary sizes required for particular text types and there may be differences in difficulty levels across forms or language versions of the same test, we were interested in how the VST performed across languages. This approach is supported by the fact that the VST is based on corpora designed and frequency lists analyzed the same way in the Routledge Frequency Dictionaries book series as well as on the same design and the same quality assurance processes of test construction and item validation as implemented by the Institute of Test Research and Test Development.

times the standard deviation (SD) was selected to establish cut scores. This measure included the vast majority of the test takers who reached a particular band (approx. 93%).

**Table 2**

*Descriptive Total Score Statistics of Each Band*

| | N | Min | Max | Median | Mean | S.E. | SD | M-1.5*SD |
|---|---|---|---|---|---|---|---|---|
| Below Band 1 | 806 | 0 | 50 | 8.50 | 12.37 | 0.43 | 12.23 | 0 |
| Band 1 | 148 | 22 | 60 | 41.5 | 40.94 | 0.73 | 8.86 | 28 |
| Band 2 | 150 | 15 | 67 | 53 | 51.72 | 0.70 | 8.56 | 39 |
| Band 3 | 58 | 44 | 71 | 61 | 60.48 | 0.75 | 5.74 | 52 |
| Band 4 | 35 | 56 | 76 | 69 | 68.77 | 0.76 | 4.48 | 62 |
| Band 5 | 44 | 55 | 86 | 75 | 74.05 | 1.04 | 6.91 | 64 |
| Total | 1241 | | | | | | | |

Many of the test takers who were rated below Band 1 had a score of 0 correct. The Productive VST gives credit only for words that are used grammatically and orthographically correctly. Especially the US test takers included many beginning language learners who may have been overwhelmed by the test. For these examinees, the last column includes the actual scores of the test takers, which could not be lower than 0, rather than the result of the equation.

Table 3 displays suggested correspondences between total scores and bands based on the descriptive total score statistics of each band. It shows the minimum score needed for a given band, assuming that the test stops after the band in question; the mean minus 1.5 times the standard deviation from Table 2; and the suggested total scores required to be assigned to a particular band.

**Table 3**

*Suggested Bands Based on Total Scores for the Productive VST*

| Minimum Score Needed | M-1.5*SD | Suggested Total Scores | Band |
|---|---|---|---|
| | | 0-27 | Below Band 1 |
| 13 | 28 | 28-38 | Band 1 |
| 26 | 39 | 39-49 | Band 2 |
| 39 | 52 | 50-58 | Band 3 |
| 52 | 62 | 59-64 | Band 4 |
| 65 | 64 | 65-90 | Band 5 |

The suggested total score for each band is slightly lower than the mean minus 1.5 SD for Bands 3 and 4 to spread cut scores more evenly at the higher bands. Note particularly the bunched scores for Bands 4 and 5 in the column M-1.5*SD. In addition, the suggested cut score for Band 5 uses the minimum score needed. Suggested total scores for Bands 3 and 4 were lowered by 2 and 3 points, respectively, from the minimum score required, while Band 5 was raised by 1 point.

To examine the relationship between total scores, bands based on 70% correct and suggested new bands, bivariate correlations between these three measures were run. Table 4 shows both *Pearson* and *Spearman* correlations. All correlations were significant at the .01 level (2-tailed).

**Table 4**

*Correlations Between Total Scores and Bands Assigned According to Total Score vs. According to the 70% Correct Criterion*

| Correlation Between Total Score and | *Pearson's R* | *Spearman's Rho* |
|---|---|---|
| Band Assigned According to Total Score | .938 | .904 |
| Band Assigned According to 70% Correct | .845 | .819 |

*Note:* $p < .01$.

Table 4 demonstrates that the correlation between total score and bands assigned according to Table 3 is markedly stronger than assignments according to the standard algorithm focusing on individual bands, because it uses more of the more fine-grained information contained in total scores. It is our contention that this may estimate the vocabulary size of a test taker slightly better than yes/no decisions based on the 70% correct cut-off.

To look for evidence of convergent validity, correlations between three measures of proficiency, i.e., listening, speaking, and reading proficiency and the two ways of defining productive vocabulary bands were examined. Proficiency was measured using the results of official ACTFL Reading Proficiency Tests, Listening Proficiency Tests, and Oral Proficiency Tests (OPIc). There were 592 RPTs, 607 LPTs, and 195 OPIc with productive vocabulary scores. Tables 5 to 7 demonstrate how many test takers were at a particular proficiency level, while Figures 1 to 3 show boxplots of the P-VST results by ACTFL proficiency level.[2]

**Table 5**

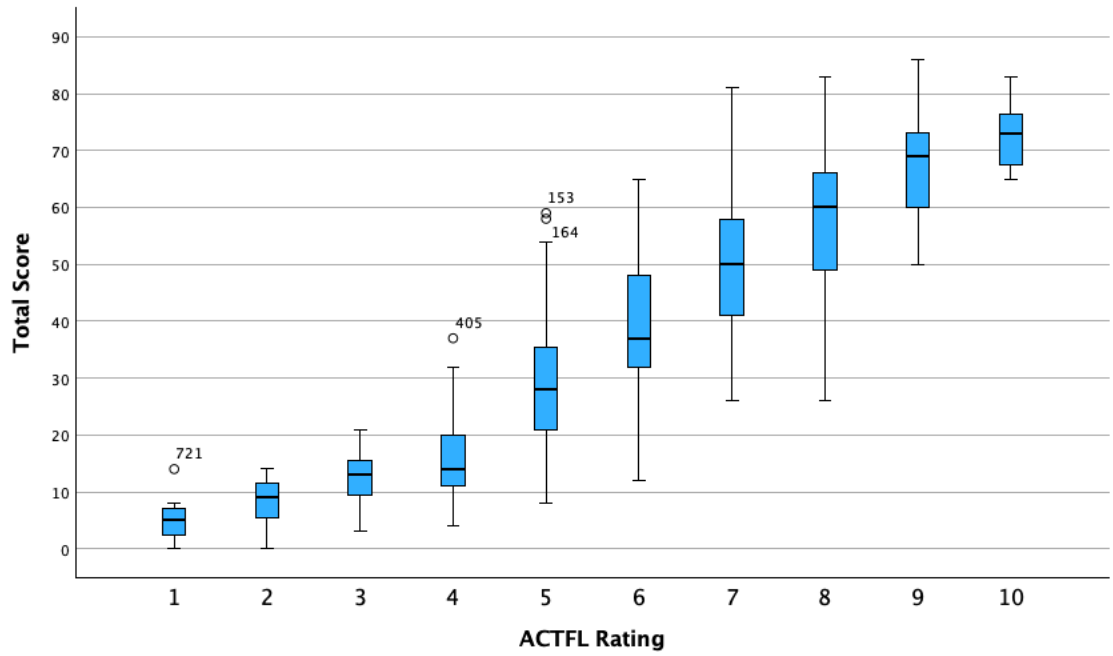
*Number of Test Takers per Reading Proficiency Level*

| ACTFL Level | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N | 8 | 12 | 20 | 65 | 83 | 81 | 169 | 114 | 29 | 11 | 592 |

[2] 1 = NL (Novice Low); 2 = NM (Novice Mid); 3 = NH (Novice High); 4 = IL (Intermediate Low); 5 = IM (Intermediate Mid); 6 = IH (Intermediate High); 7 = AL (Advanced Low); 8 = AM (Advanced Mid); 9 = AH (Advanced High); 10 = S (Superior). NH and IL are roughly equivalent to A1; IM to A2; IH and AL to B1; AM to B2; AH and S to C1 (ACTFL, n.d.).

**Figure 1**

*Boxplots of P-VST Scores by ACTFL Reading Proficiency Rating*




*Note:* 1 = NL; 2 = NM; 3 = NH; 4 = IL; 5 = IM; 6 = IH; 7 = AL; 8 = AM; 9 = AH; 10 = S.

Figure 1 demonstrates that total scores increase as proficiency levels increase. Total scores increase modestly at levels NL (1) to IL (4) and more steeply from IL (4) to AH (9), while somewhat leveling off from AH (9) to S (10). It also shows that Advanced High (AH) and Superior (S) readers achieved very high overall scores.

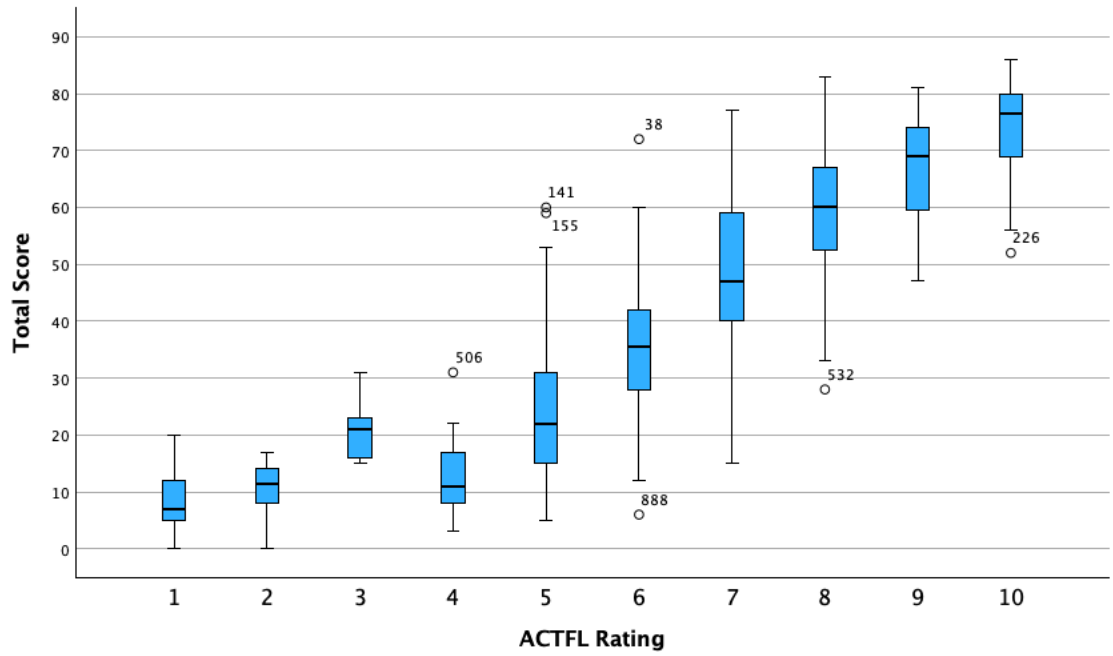
**Table 6**

*Number of Test Takers per Listening Proficiency Level*

| ACTFL Level | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N | 20 | 6 | 5 | 34 | 106 | 90 | 193 | 123 | 20 | 10 | 607 |

**Figure 2**

*Boxplots of P-VST Scores by ACTFL Listening Proficiency Rating*




*Note:* 1 = NL; 2 = NM; 3 = NH; 4 = IL; 5 = IM; 6 = IH; 7 = AL; 8 = AM; 9 = AH; 10 = S**.**

Figure 2 demonstrates that, in general, total scores increase as proficiency levels increase. Total scores increase very modestly from levels NL (1) to IL (4) and more steeply from IL (4) to S (10). The anomaly is NH (3), which consisted only of 5 test takers and may be an artifact of the present population. Figure 2 also shows that Advanced High (9) and Superior (10) readers achieved very high overall scores.

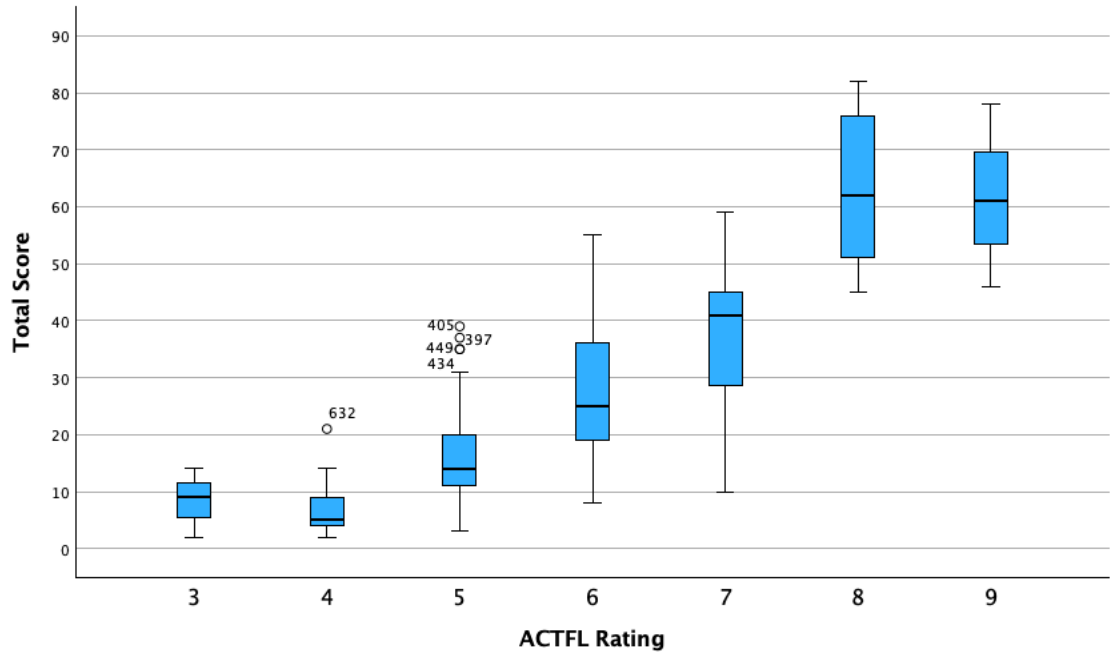
**Table 7**

*Number of Test Takers per Speaking Proficiency Level*

| ACTFL Level | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N | 0 | 0 | 3 | 25 | 101 | 41 | 16 | 6 | 3 | 0 | 195 |

**Figure 3**

*Boxplots of P-VST Scores by ACTFL Oral Proficiency Rating*




*Note:* 1 = NL; 2 = NM; 3 = NH; 4 = IL; 5 = IM; 6 = IH; 7 = AL; 8 = AM; 9 = AH; 10 = S.

Figure 3 shows that, in general, total scores increase as proficiency levels increase. There were only 3 test takers whose speaking proficiency was NH (3) and there were no test takers with proficiency levels below that. There were noticeable increases in word knowledge from IL (4) to AM (8), while the highest levels again only consisted of a handful of test takers, 6 at AM (8) and 3 at AH (9), making these results less generalizable. Figure 3 also shows that advanced Mid (8) and high (9) speakers achieved high overall scores.

Table 8 shows both *Pearson* and *Spearman* correlations between the three skills and the two ways of calculating vocabulary bands. All correlations were significant at the .01 level (2-tailed).

**Table 8**

*Correlations Between Bands Assigned According to Total Score and 70% Correct per Band and Reading, Listening, and Speaking Proficiency Levels*

| | Total Score Bands | | | 70% Bands | |
|---|---|---|---|---|---|
| | N | *Pearson's R* | *Spearman's Rho* | *Pearson's R* | *Spearman's Rho* |
| Reading | 592 | .771 | .811 | .666 | .734 |
| Listening | 607 | .717 | .794 | .633 | .732 |
| Speaking | 195 | .716 | .609 | .603 | .478 |

*Note:* $p < .01$.

Table 8 demonstrates that correlations between bands and reading, listening, and speaking proficiency levels are considerably higher for bands derived from total scores than bands defined by 70% correct for all three skills. This provides strong convergent validity evidence to suggest that the classification of bands according to total score may be more useful than the 70% correct model. In addition, it provides evidence that the total score definitions of each band are valid.

To further examine the relationship between vocabulary and reading, Tables 9 and 10 show crosstabulations of vocabulary bands and ACTFL reading proficiency levels for the two methods of assigning bands. Cells that include at least 10% of the total number of the respective band are highlighted.

**Table 9**

*Crosstabulation of ACTFL Reading Proficiency Levels and Vocabulary Bands Assigned According to Total Score*

| Band | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 8 | 12 | 20 | 59 | 41 | 12 | 4 | 1 | 0 | 0 | 157 |
| 1 | 0 | 0 | 0 | 6 | 25 | 33 | 31 | 5 | 0 | 0 | 100 |
| 2 | 0 | 0 | 0 | 0 | 12 | 24 | 48 | 24 | 0 | 0 | 108 |
| 3 | 0 | 0 | 0 | 0 | 4 | 8 | 45 | 20 | 5 | 0 | 82 |
| 4 | 0 | 0 | 0 | 0 | 1 | 3 | 24 | 26 | 6 | 0 | 60 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1 | 17 | 38 | 18 | 11 | 85 |
| Total | 8 | 12 | 20 | 65 | 83 | 81 | 169 | 114 | 29 | 11 | 592 |

*Note: Cells that include at least 10% of the total number of the band are highlighted.*

**Table 10**

*Crosstabulation of ACTFL Reading Proficiency Levels and Vocabulary Bands Assigned According to 70% Correct of Individual Bands*

| Band | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 8 | 12 | 20 | 62 | 65 | 36 | 28 | 10 | 0 | 0 | 241 |
| 1 | 0 | 0 | 0 | 3 | 10 | 27 | 46 | 16 | 2 | 0 | 104 |
| 2 | 0 | 0 | 0 | 0 | 7 | 15 | 60 | 41 | 8 | 0 | 131 |
| 3 | 0 | 0 | 0 | 0 | 1 | 2 | 19 | 17 | 3 | 3 | 45 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 14 | 6 | 4 | 33 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1 | 7 | 16 | 10 | 4 | 38 |
| Total | 8 | 12 | 20 | 65 | 83 | 81 | 169 | 114 | 29 | 11 | 592 |

*Note: Cells that include at least 10% of the total number of the band are highlighted.*

Tables 9 and 10 demonstrate that a scoring system that focuses on yes/no decisions per band (Table 10) appears to make it more difficult to receive a higher vocabulary score. 241 test takers, e.g., received a score of Below 1000 in Table 10, whereas 157 did in Table 9. Table 10 also shows a surprisingly large number of IH and AL readers with a vocabulary score of less than 1,000

words in addition to a large number of AM readers with very low vocabulary scores, between less than 1,000 and 2,000 words. In general, however, both tables show a clear association of higher vocabulary scores with higher proficiency levels.

Tables 11 and 12 show crosstabulations of vocabulary bands and ACTFL listening proficiency levels for the two methods of assigning bands. Cells that include at least 10% of the total number of the respective band are highlighted.

**Table 11**

*Crosstabulation of ACTFL Listening Proficiency Levels and Vocabulary Bands Assigned According to Total Score*

| Band | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 20 | 6 | 4 | 33 | 70 | 20 | 5 | 0 | 0 | 0 | 158 |
| 1 | 0 | 0 | 1 | 1 | 28 | 36 | 39 | 7 | 0 | 0 | 112 |
| 2 | 0 | 0 | 0 | 0 | 4 | 20 | 64 | 18 | 1 | 0 | 107 |
| 3 | 0 | 0 | 0 | 0 | 2 | 11 | 35 | 30 | 3 | 2 | 83 |
| 4 | 0 | 0 | 0 | 0 | 2 | 2 | 29 | 24 | 5 | 0 | 62 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1 | 21 | 44 | 11 | 8 | 85 |
| Total | 20 | 6 | 5 | 34 | 106 | 90 | 193 | 123 | 20 | 10 | 607 |

*Note: Cells that include at least 10% of the total number of the band are highlighted.*

**Table 12**

*Crosstabulation of ACTFL Listening Proficiency Levels and Vocabulary Bands Assigned According to 70% Correct of Individual Bands*

| Band | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 20 | 6 | 5 | 33 | 90 | 48 | 42 | 6 | 0 | 0 | 250 |
| 1 | 0 | 0 | 0 | 1 | 12 | 25 | 52 | 16 | 1 | 0 | 107 |
| 2 | 0 | 0 | 0 | 0 | 3 | 14 | 64 | 44 | 5 | 2 | 132 |
| 3 | 0 | 0 | 0 | 0 | 1 | 2 | 16 | 24 | 4 | 0 | 47 |
| 4 | 0 | 0 | 0 | 0 | 0 | 1 | 10 | 18 | 3 | 1 | 33 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 15 | 7 | 7 | 38 |
| Total | 20 | 6 | 5 | 34 | 106 | 90 | 193 | 123 | 20 | 10 | 607 |

*Note: Cells that include at least 10% of the total number of the band are highlighted.*

Tables 11 and 12 also demonstrate that a scoring system that focuses on yes/no decisions per band (Table 12) makes it more difficult to receive a higher vocabulary score. 250 test takers, e.g., received a score of Below 1000 in Table 12, whereas 158 did in Table 11. Table 12 also shows a large number of IH and AL listeners with a vocabulary score of less than 1,000 words, which is an unlikely combination. In general, however, both tables show a clear association of higher vocabulary scores with higher proficiency levels.

Tables 13 and 14 show crosstabulations of vocabulary bands and ACTFL speaking proficiency levels for the two methods of assigning bands. Cells that include at least 10% of the total number of the respective band are highlighted.

**Table 13**

*Crosstabulation of ACTFL Speaking Proficiency Levels and Vocabulary Bands Assigned According to Total Score*

| Band | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 3 | 25 | 89 | 22 | 4 | 0 | 0 | 0 | 143 |
| 1 | 0 | 0 | 0 | 0 | 11 | 10 | 2 | 0 | 0 | 0 | 23 |
| 2 | 0 | 0 | 0 | 0 | 1 | 7 | 8 | 1 | 1 | 0 | 18 |
| 3 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 0 | 0 | 4 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 0 | 4 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 3 |
| Total | 0 | 0 | 3 | 25 | 89 | 22 | 4 | 0 | 0 | 0 | 143 |

*Note: Cells that include at least 10% of the total number of the band are highlighted.*

**Table 14**

*Crosstabulation of ACTFL Speaking Proficiency Levels and Vocabulary Bands Assigned According to 70% Correct of Individual Bands*

| Band | NL | NM | NH | IL | IM | IH | AL | AM | AH | S | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 3 | 25 | 94 | 32 | 9 | 0 | 0 | 0 | 163 |
| 1 | 0 | 0 | 0 | 0 | 7 | 3 | 3 | 1 | 1 | 0 | 15 |
| 2 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 2 | 0 | 0 | 7 |
| 3 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 1 | 0 | 0 | 6 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 4 |
| Total | 0 | 0 | 3 | 25 | 101 | 41 | 16 | 6 | 3 | 0 | 195 |

*Note: Cells that include at least 10% of the total number of the band are highlighted.*

Tables 13 and 14 again demonstrate that a scoring system that focuses on yes/no decisions per band (Table 14) makes it more difficult to receive a higher vocabulary score. 163 test takers, e.g., received a score of *Below 1000* in Table 14, whereas 143 did in Table 13. Table 14 also shows a larger number of IH and AL speakers with a vocabulary score of less than 1,000 words than Table 13.

Moreover, Table 14 has a break at Band 4. There are no oral proficiency scores associated with it. In general, however, both tables show a clear association of higher vocabulary scores with higher proficiency levels.

On the basis of the preceding discussion, Table 15 presents the suggested total score cut scores for frequency bands of the most frequent 5,000 words of a language as measured by the ITT P-VST. Note that these cut scores are the same as in Table 3.

**Table 15**

*Total Score Cut Scores for Frequency Bands*

| Total Score | 0-27 | 28-38 | 39-49 | 50-58 | 59-64 | 65-90 |
|---|---|---|---|---|---|---|
| Bands | Below 1000 | 1000 | 2000 | 3000 | 4000 | 5000 |

## Conclusion

The present analysis based on 1,241 examinees taking the Productive Vocabulary Size Test (P-VST) provides solid evidence for the argument that frequency band definitions based on the total score of examinees are as valid and reliable as the established model of rating each band separately and may be more useful because they can be assigned more rapidly and they use more of the information provided by the test. In addition, the present study provides solid validity evidence of the cut scores established by the analysis.

There is one caveat, however. Assigning bands according to total score using the formula above assumes test takers having completed all five bands. In case of doubt, it is suggested to calculate bands using both methods of assigning bands and to allocate the higher band.

**References**

ACTFL (n.d.). *Assigning CEFR Ratings to ACTFL Assessments*. https://www.actfl.org/assessments/assigning-cefr-ratings-to-actfl-assessments

Institute for Test Research and Test Development (n.d.). Vocabulary Tests. https://itt-leipzig.de/about-the-vocabulary-tests-2/?lang=en

Möhring, J. (2022). Sprachliche Studierfähigkeit im Spiegel produktiver und rezeptiver Wortschatzkompetenz. *Informationen Deutsch als Fremdsprache*, 49(4), 384-410. https://doi.org/10.1515/infodaf-2022-0058

Spino-Seijas, L., & Tschirner, E. (2023, August 29). *URI Modern Languages Proficiency Initiative: Spring 2023 Assessment Cycle* [Conference presentation]. Modern and Classical Languages and Literatures Retreat, University of Rhode Island, Kingston, RI, United States.

Tschirner, E. (2021). *Examining the validity and reliability of the ITT vocabulary size tests. Research Papers in Assessment: Vol. 3.* Universitätsbibliothek Leipzig. https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-763177

Tschirner, E. & Möhring, J. (2024). *Examining the validity and reliability of the Productive Vocabulary Size Test*. *ITT Technical Reports on Language Testing: Vol. 2*. Universitätsbibliothek Leipzig. https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-916433

Wei, X., Gunter, T. C., Adamson, H., Schwendemann, M., Friederici, A. D., Goucha, T., & Anwander, A. (2023). *White matter plasticity during second language learning within and across hemispheres*. bioRxiv. https://doi.org/10.1101/2023.04.21.537810