



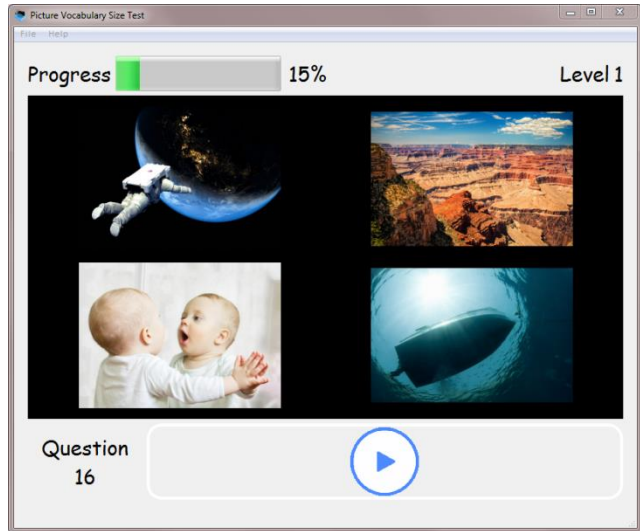
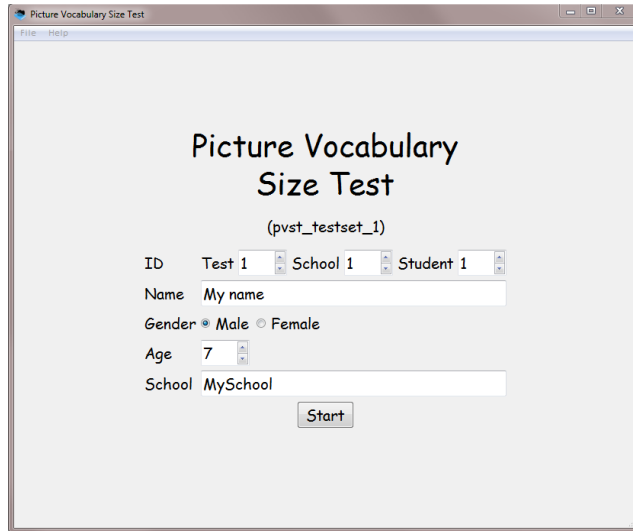
Picture Vocabulary Size Test (PVST)

Build 1.1.0 (Released July 13, 2017)

¹Laurence ANTHONY and ²Paul NATION

¹Center for English Language Education in Science and Engineering, Faculty of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.

²School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand. Help file version: 001.



Introduction

The *Picture Vocabulary Size Test (PVST)* is a test of receptive vocabulary size. The test measures whether the test-taker can find a suitable meaning (a picture) for a given partly contextualized word form. It is a recognition test primarily intended for young pre-literate native speakers up to eight years old and young non-native speakers of English. Two 96-item test sets are included with the test. These test sets use different questions but following the same design procedures. The PVST was designed by Paul Nation of Victoria University of Wellington, New Zealand and implemented as a software package by Laurence Anthony of Waseda University, Japan. Jannie van Hees of The University of Auckland, New Zealand played an important role in trialing the test.

The *PVST* requires a configuration file (`config.txt`), a scripts file (`scripts.txt`) in a "scripts" folder, a set of associated audio files in an "audio" folder, and a set of image files in a "pictures" folder. The configuration file and the scripts file should be **UTF-8 encoded**.

The *PVST* runs on any standard computer, touch-screen computer, or tablet computer running Microsoft Windows (tested on Win 7), Macintosh OS X (tested on OS X 10.9 Mavericks), or Linux (tested on Linux Mint 17) computers. It is developed in Python and Qt using the *PyInstaller* compiler to generate executables for the different operating systems.

Getting Started (No installation necessary)

Windows

On Windows systems, simply double click the *PVST* icon to launch the program.

Macintosh OS X

On Macintosh systems, simply double click the *PVST* zip file. The zip file will unzip the *PVST* application. Throw away the zip file once the *PVST* application has been unpacked.



Linux

On Linux systems, set the permissions to run the executable, then double click the *PVST* icon to launch the program.

Aims and target test takers of the PVST

The PVST is intended to be used for formative assessment, particularly diagnostic assessment, where performance on a task such as reading or participating in content-focused oral discussion may be affected by vocabulary size. The PVST is a largely decontextualized measure of receptive vocabulary knowledge to some degree isolating vocabulary knowledge from other skills. As a result, a learner's performance on the PVST can be used as an indication of whether vocabulary knowledge is likely to be a major factor accounting for poor performance on a task. Thus, the test can be used along with other diagnostic procedures for examining a learner's reading, listening, speaking, and writing skills.

The PVST is primarily intended for use with young pre-literate children from the age of five onwards. With this age group, the test administration needs to take account of the fact that children are experiencing cognitive, social, emotional, and physical growth, are learning literacy skills, and are vulnerable (McKay, 2006). In essence, this means that the test needs to be administered one-on-one in a strongly supportive and flexible way to take account of the individual learner. For example, it is recommended that the administrator sit next to the child, explain how to do the test, and keep the child motivated and on task by providing encouraging comments such as "You are doing well" and "Good" throughout the test. The design of the test also takes account of the young age of its target users with the use of colourful pictures, oral and written cues, a multiple-choice format, and a possible touch-screen-based delivery involving pointing and touching.

Although the test is designed primarily for young pre-literate children, it can be used with older learners of English as a second or foreign language, especially those with literacy problems. However, some test items will be more likely known by children so the results need to be interpreted with care.

Administering the PVST

The PVST can be administered through the following steps:

- Step 1: The administrator loads a set of test items for use with the test via the File menu "Load test items" option. Test item sets are available at <http://www.laurenceanthony.net/software/pvst/tests/>. Any test items sets placed in a folder labeled "tests" together with the main software executable will be listed automatically in File menu. The first test set will be loaded automatically.
- Step 2: The administrator initializes the test by inputting the learner's biographic information (name, age, gender, school,) and pressing "Start".
- Step 3: The learner proceeds through the test by pressing on each "Listen" icon, hearing and reading each question prompt, and selecting a matching picture (or choosing a "Don't Know" button if the option is activated in the test settings). A red line appears around the selected picture. With the default settings, the test moves automatically to the next screen. Setting the config.txt file "autoChange" option to FALSE, causes the program to add an arrow to the interface, which the learner must press to proceed to the next screen.
- Step 4: The learner can listen to a question again by re-touching the speaker icon twice. The learner can change their answer choice by selecting a different picture (or choosing a "Don't Know" button if the option is activated in the test settings).

Step 5: On completing the test, the learner's estimated vocabulary level is shown and the data is stored in the "results" folder. Each learner's individual test results are stored in a plain text file where the first digit in the file name records the test ID, the second two digits record the school ID, and the final three digits record the learner ID (e.g. "test_taker_101001.txt"). The test responses and response times for all learners are stored in a single file called "complete_results.txt". When interpreting the results, one correct answer corresponds to a vocabulary size of 62.5 words.

Example: Score 54/96 → 54 x 62.5 → 3,375 vocabulary size

The test normally takes around 15 minutes to complete. However, there is no time limit for the test. Experience shows that some learners may take as long as 25 minutes, agonizing over answer choices and giving it their best shot.

When administering the test, it is useful to make sure that all learners are aware of the following.

- They know how to initiate a test item by pressing the audio button.
- They should look at all four pictures when considering choices.
- They know they can press the audio button again to hear a sentence again.
- They know how to select an answer option by clicking (or touching) one of the pictures.
- They can change their mind about a choice, by clicking (or touching) a different picture.
- They can press the "I don't know" button (if activated) if they have no idea of the correct option.

Each of the above points should be explained to each learner with a demonstration, perhaps with the use of a short checklist. Also, the administrator should be watching each learner to ensure they perform each procedure when needed. Where learners respond too quickly, the administrator should ask them if they looked at all four pictures, encourage them to do this, and praise them for doing it.

Administrators of the PVST need to be aware of the following issues:

- Learners can adopt very different strategies when taking the test. Some older learners (e.g. eight year olds) may carefully re-listen to sentences and consider all choices before selecting the correct answer. Other learners may want to rush through the test somewhat in the manner of a video game. In this case, the test administrator should restrain the learners and encourage them to consider all pictures and to think carefully about their choices.
- Some learners may deliberately avoid choosing the "I don't know" option (if activated), insisting that they should try to answer every test item. Other learners may be happy to indicate their lack of knowledge.
- Using a touch screen computer raises hygiene issues. Some young learners may put their fingers in their mouths or up their noses, touch the screen, and then put their fingers back in their mouth or up their noses. It is recommended that any device is thoroughly cleaned after each iteration of the test.
- Learners often appear very happy to take the test, and a few may spontaneously speak aloud about their thinking processes as they complete test items. Care should be taken that the learners do not get overly excited about the test and stop focusing on the test items.

Interpreting the PVST

The PVST is based on the most frequent 6000 word families of English for young native-speaking children. Each test word in the PVST represents 62.5 words in the source lists. So, a learner's score on the test needs to be multiplied by 62.5 to get their total vocabulary size. Thus, a learner with a score of 54 has an approximate receptive vocabulary size of 3,375 word families. The program makes this calculation in the results it produces. It is important to note that the test is a sensitive receptive recognition test, and thus, it gives credit for partial knowledge. This means that getting a word correct on the test does not necessarily mean that the word is well known.

As only the most frequent 6000 word families appear in the PVST, if the test is used with native-speakers of English who are more than eight years old, a ceiling effect on the test scores may be observed as older learners are likely to have a substantially larger vocabulary size than that shown in the test. Native-speaking learners should sit the whole test because even five year olds may know words from the 5th and 6th 1000 levels.

The rough rule of thumb for a native-speaker's vocabulary size is to take 2 away from the learner's age and multiply the number by 1000 to find their likely vocabulary size. So, a seven year old is likely to have a receptive vocabulary size of around 5000 word families. This rule of thumb fits with Biemiller's (2005) data from young Canadian children and from data gathered in the trialling of this test. There can be wide variation of vocabulary size at any particular age level.

The safest interpretation of the test results is to look for positive results. That is, the results should be used in order to exclude vocabulary size as a factor negatively affecting past performance in, for example, reading. This kind of interpretation assumes that there are minimum vocabulary sizes sufficient for the performance of certain tasks at various age levels. For example, a child with a receptive vocabulary size of 4000 word families knows enough vocabulary to learn to read English without vocabulary size being a major issue negatively affecting this learning. Of course, the child may not be able to recognize the written words that they know in a spoken form, but that is a different problem to vocabulary size.

It is very important that poor performance on the test is carefully investigated to make sure that the result truly represents the child's knowledge. At the very least, poor performance on the test (very low scores) should be checked by administering the other version of the test at a different time under the most favourable supportive circumstances, perhaps getting the learner to talk aloud while taking the test. Poor results on the PVST should also lead to a further investigation of a particular learner's vocabulary size or an investigation of other aspects that may affect performance. Factors may include:

- the administration of the test
- bias in the test through the pictures used or the computerised presentation
- the learner's vision and hearing
- the learner's feelings about being tested
- the learner's first language (i.e., is English the learner's first, second, or additional language)
- the length of time the learner has been learning English if it is not their first language

A close analysis of results on the PVST can be carried out using the *logger.txt* file which is created in the results folder and lists each learner's result for each word. By looking at this file, test administrators can see which words are known by everyone who sat the test, which are known by some of the learners, and which words nobody knows. The table below shows part of the *logger.txt* file. The first number is each learner's unique id number created by the program. This is followed by their name, age, name of school, total score, and total score converted to a vocabulary size (by multiplying by 62.5). The following numbers (0 or 1) are the results for each test item, with 0 indicating an incorrect or skipped answer and 1 indicating a correct answer. There are five of these columns in the example but in the full *logger.txt* file, there are 96 – one for each test item.

101001	Janet	8	Myschool	59	3687	0	0	0	1	1
101002	John	7	Myschool	67	4187	1	0	1	1	1
101003	Bill	7	Myschool	64	4000	1	0	0	0	1
101004	Janet	7	Myschool	64	4000	1	0	0	0	1

Trouble shooting

If the PVST does not launch correctly, it is important to check that all files and folders needed by the program are correctly stored in the target test folder. The files and folders needed by PVST are as follows:

- *scripts*

- This folder stores a UTF-8 encoded text file (*scripts.txt*) that supplies the written text for the program. It contains 96 lines
- *audio*
 - This folder contains the recordings for the audio for the text – 96 .wav files)
- *pictures*
 - This folder contains the pictures used in the test – 384 .jpg files).
- *results*
 - This folder is automatically created if not already available. It stores the results of the test.
- *config.txt [optional]*
 - This is a UTF-8 encoded text file that sets the main properties of the test. Editing this file allows a test administrator change the settings and format of the test. If this file is unavailable, an internal configuration file will be used. The config.txt file contains comments explaining the different parameters.

All new editions and bug fixes to the PVST are listed in the revision history below.

THE DESIGN OF THE TEST

Test construct

The PVST is a discrete, selective, relatively context-independent vocabulary test presented in a multiple-choice format. The following description of the test construct is based on Read and Chapelle's (2001) framework. Test-takers are required to select the best picture for each word from four choices (or select an "I don't know" option if this is activated in the test). The PVST is designed to measure both first language and second language learners' spoken receptive vocabulary size in English. The test measures knowledge of the spoken word form, the form-meaning connection, and to a smaller degree concept knowledge. The four choices are not closely related in meaning to each other. This makes it possible for a learner to answer test items in the test with a partial knowledge of the words involved.

Inferences: Although the test items in the PVST are presented in simple non-defining contexts, the test is essentially following a trait-definition of vocabulary which means that vocabulary knowledge is tested independently from contexts of use. At the item level, the test measures receptive knowledge of a word form. At the test level it provides an estimate of total vocabulary size where vocabulary knowledge is considered as including only single words (not multi-word units), and vocabulary size does not include proper nouns, transparent compounds, marginal words (e.g. *um, er, gee, gosh*), and abbreviations. The test does not measure the ability to distinguish homonyms and homographs. A ceiling effect may be observed on the test for some learners as the test only goes up to the 6th 1000 word family level. According to Biemiller (2005: 226), this may be enough for native-speaking seven to eight year olds.

The PVST does not measure skill in vocabulary use, such as productive spoken use or receptive use in reading. That is, learners may know particular words enough to correctly select what they mean, but may not know them well enough to use them either receptively or productively. If a learner correctly identifies the picture matching a test item in the test, then their learning needs to focus on using the word rather than on the word itself.

Uses: For instructional purposes, the results of the PVST can be used to guide syllabus design, choose texts for extensive reading, and help in vocabulary instruction. For research purposes, the test can be used as a measure of total receptive spoken vocabulary size for young native-speakers below the age of eight and young intermediate proficiency non-native speakers.

Impacts: If the PVST is used as intended, it is a relatively low stakes test for learners. One consequence is that it might underestimate the vocabulary size of learners who are not motivated to perform to the best of their

ability, resulting in faulty instructional decisions being made about their vocabulary learning needs. To overcome this limitation, the test should be administered to students on a one-to-one basis with the administrator present to keep the learner on task.

Test format

The test contains 96 test items, each presented in a four-choice multiple-choice format. Each test word appears in a non-defining sentence where the context uses only words from the first 500 word list. The context provides a small degree of orientation towards the meaning of the word for the learners, in that it indicates a particular sense of the word and its part of speech. The sentences are single clause sentences that are less than six words long. Although the first 500 word-families are tested, it is assumed that the learners sitting the test know the words which make up the context sentences.

The multiple-choice format for testing vocabulary knowledge has received growing criticism for two major reasons. Firstly, by providing choices rather than requiring the learner to supply a response, it does not closely represent the kind of vocabulary knowledge needed for listening or reading. In designing the PVST, this disadvantage was weighed against the practicality of using a multiple-choice format. However, it deserves further research. Secondly, the use of choices allows for the possibility of getting test items correct through guessing which may inflate scores.

Multiple-choice vocabulary questions can be answered using a range of strengths of word knowledge. This range goes from only being able to select a meaning from a small number of choices not closely related to each other to having a strong productive knowledge of a word. This means that for some words answered correctly, the learners may be at an early stage of developing knowledge of these words. Unfortunately, the data from one test format will not be enough to show which particular words are well-known, which are moderately well-known, and which are only partly known. Test users need to be aware that although every correct answer on the test is given a 1-point score, a learner's knowledge of each word is not necessarily of equal strength.

The "*I don't know*" option in the test should only be activated after careful consideration. Initially, the option may appear to reduce the possibility of learners guessing the correct answer. However, research on adults has been shown it to be a complicating factor in multiple-choice vocabulary tests, with some test takers choosing to use the option frequently and others never using it (Zhang, 2013; Stoeckel, Bennett & McLean, 2016).

Test creation

The source corpus and lists

Twelve 500 word lists were created from a corpus of 5 million tokens. One million tokens consisted of books used with learners just beginning to read, and texts from the *New Zealand School Journals* Parts One, Two, Three, and Four. This sub-corpus was used exclusively for the first three 500 word lists (1,500 word families). The remainder of the corpus was one million tokens of movies and TV programs, two million tokens of colloquial spoken American English, and one million tokens of spoken British English. A strength of the test is that it is based on lists derived from a corpus at least partially suited to young learners. The data from the corpus was modified using data from Biemiller's (2010) analysis of *The Living Word Vocabulary* (Dale & O'Rourke, 1981). Data from the Oxford Children's Corpus (Banerji et al, 2012) was used to check the lists and develop the two 500 word lists at the 6000 word level (lists 11 and 12). The words in the first 500 word list are typically more frequent and more likely to be known than words in the second 500 word list and so on. A weakness of the lists is that they use Level 6 word families (Bauer & Nation, 1993). It would have been better to use lists based on Level 3 partial families which are now available, but they were not available when the test was designed.

The first 3000 words of the BNC/COCA lists are almost all in the children's lists (all of the BNC/COCA first 2000 and 918 of the BNC/COCA 3rd 1000). However, the first 1000 words of the BNC/COCA lists are not the same as the first two 500-word children's lists, although there is an overlap of 811 word families. The children's lists

contain words from every one of the first twenty five word family levels of the BNC/COCA lists. The table below shows some example words from the children's lists with their BNC/COCA list levels in brackets. The word *biff*, for example, is at the 14th 1000 level in the BNC/COCA lists, and *budgie* is at the 15th 1000 level. Native-speaking children tend to know these words.

biff (14)	bazooka (16)	gobsmacked (18)	abracadabra (20)	diplodocus (22)
bumblebee (14)	cheep (16)	hotdog (18)	boing (20)	gobstopper (23)
budgie (15)	earwig (17)	icky (19)	gymkhana (21)	brachiosaur (24)
ladybird (15)	humongous (17)	triceratops (19)	jurassic (21)	passionfruit (24)

The first 500 word list contains all numbers to the thousands, words signalling specific times (e.g., *today*, *yesterday*), greetings (e.g. *hello*, *goodbye*), expressions of gratitude (e.g. *thanks*), meals (i.e., *breakfast*, *lunch*, *dinner*), seasons (i.e., *spring*, *summer*, *autumn*, *winter*), and parts of the day (i.e., *morning*, *afternoon*, *evening*, *night*). The fourth 500 word list (1,500-2,000) contains all the days of the week and the months of the year.

Sampling question items from the lists

A sample of eight target words was selected from each 500 word-family level (a 1 in 62.5 sample) to form the complete 96-item test. The twelve 500 word lists were each sorted by their frequency in the 5-million word corpus. Then, two sets of test items were sampled from the results of the sorting. Sample 1 in each 500 word block began with a frequency-ranked item 10, then item 70, 130, 190, 250, 310, 370, 450 and if necessary 490, from the 500 items in the level. The exception was the first 500 level where the sampling began with item 70 and continued with every 60th item to item 490. Sample 2 similarly chose every 60th word beginning with frequency ranked item 5 then 65, 125, 185, 245, 305, 365, 425 the exception was the first 500 level where the sampling began with item 65 and continued with every 60th item to item 485 this method of sampling was used to ensure that a representative range of frequencies within each 500 word level were in each 500 word level. Both function words and content words were chosen.

In sample 1, the word *breath* was chosen in the 2nd 500 block and *breathe* in the 3rd 500 block. To avoid having almost identical words appearing twice in the test, the word *earn*, which immediately preceded *breathe* was chosen instead.

Choosing distractors

The distractors were chosen from the 500 word-family levels in a similar way to the test words. However, as the part of speech of the distractor had to match that of the test item, it was sometimes necessary to choose a different word from the level than the one at the specific sampling point.

Choosing pictures

The pictures for the tests were downloaded from the website www.shutterstock.com using two one-month licenses which allowed 25 pictures to be downloaded each day (starting 10 February 2014, and starting 14 March 2015). A very small number of pictures were bought from www.dreamtime.com. In a revision of the test (June 2014), additional pictures were downloaded from Shutterstock. Almost all pictures used in the test are photographs. Where test items are homographs, the most frequent meaning was chosen, guided by Kevin Parent's (2012) research.

The criteria used to select pictures for the test were as follows (in order of importance):

- 1 The picture most clearly represented the meaning of the word
- 2 Pictures containing young children and people from a range of racial groups were given preference
- 3 Photographs were preferred over diagrams or cartoons.

Test trialing

The two versions of the test were initially trialed with adults to make sure they could get all the test items correct. Then, the test was trialed with a ten year old child and a seven year old child, who experienced no difficulty with the test format and only had difficulty with a small number of the pictures (*shabby*, *conquest*, *battered*) which were then changed. The two children were asked about several of the test items as they sat the test to check that correct and incorrect answers reflected their knowledge of the words and not misleading aspects of the pictures or prompts.

Several changes were made to the test as a part of its development to improve the test's validity. These changes are detailed below.

- Some prompts were changed if they contained a 'give-away' word which was not the test word. For example, the prompt for the test item "chap" was changed from "He is a good chap" to "This is a good chap" as the correct picture was the only picture in the set of four that contained a person. As a result, the word "He" gave away the answer. Similarly, the prompt "It's a large flock" was replaced by "It's a flock", as "large" was helping to eliminate possible choices. Similarly, "He's working independently" was replaced by "He's doing it independently", as "working" helped eliminate choices.
- Some distractors were changed if they could have been correct answers. For example, one of the distractors for the test item "*crimson*" had to be changed as part of the distractor picture was coloured crimson.
- One or two test items that might upset young children (e.g., *nightmare*, *demon*) were replaced.
- A few pictures were replaced with those that closer matched what the children knew of the word (i.e., *limb* (tree to leg), *function* (vending machine to computer), *coordinate* (from a meeting to two children jumping), *shabby* (from a young man dishevelled with a tie to an old man with rough clothes).
- Several words in the first sample set could not be satisfactorily pictured and were replaced (i.e., *selection* was replaced with *electric*, *resume* was replaced with *thistle*, and *mature* was replaced with *objected*).
- An early version of the test using a multiple-item matching format with substantially the same items and pictures as those here was used to gather data from 71 young native speakers. The results are reported in Anthony, van Hees and Nation (forthcoming).

Test weaknesses

The format of the test leads to several unavoidable weaknesses.

- 1 Some test items do not require detailed knowledge of the target word, that is, they require just enough knowledge to distinguish it from the distractors (e.g. *savage*).
- 2 Some test items require a greater degree of world knowledge than other items. For example, the item testing *flit* requires the learner to work out that a butterfly moves in this way, and the item testing *agenda* requires that the learner understands that a meeting may involve an agenda. Items like *table* or *grass* require no such inferencing. Surprisingly, even five year olds seem to have no great difficulty performing this inferencing.
- 3 A small number of the prompts contain additional content words to the test item and so understanding the item at least partly depends on understanding the words in the context. For example, *see* is a context word in the prompt "*She saw the horizon*" and *full* is a context word in the prompt "*full of anguish*". To reduce this problem, all context words are in the first 500 words of English.
- 4 Learners sitting multiple-choice tests can benefit from applying test-taking strategies. One of the most common strategies is to eliminate the alternatives, resulting in fewer choices to focus on. However, such a test-taking strategy does not necessarily undermine the validity of the test, because it indicates close engagement with the test and a thoughtful search for the answer.
- 5 The multiple-choice item format allows for guessing, which can inflate scores. For more on this point, see the earlier discussion.
- 6 The unit of counting when making the word lists was based on Level 6 word-family rules (Bauer & Nation, 1993). This could result in the test underestimating vocabulary size because high-frequency

derived forms which young learners might treat as different words from their stem form were not tested.

- 7 Adults sitting the test inevitably criticise some of the choices of pictures even though they can see which answer is correct. In the case of doubt about a test item, it is worth checking to see if young learners have a problem with the choices as the test has been designed for them and was trialled with them.

RESEARCH QUESTIONS

The following are possible research questions that the PVST might address

Testing vocabulary size and growth

- How much vocabulary do young native speakers know when they begin primary school?
- Do young children have a large enough vocabulary size to begin learning to read?
- Do children who live in homes where more than one language is spoken have vocabulary sizes comparable to those of children who live in monolingual English-speaking homes?
- At what rate does a native speaker's vocabulary size grow?
- Do learners of the same age from different socioeconomic groups have comparable vocabulary sizes?
- At what rate does a young immigrant to a country pick up English vocabulary?
- How wide is the range of vocabulary sizes at each age level? What are the likely causes of this variation?

Testing the test

- Do the scores at each of the twelve levels of the PVST drop from one level to the next? That is, does the test show implicational scaling?
- Are the two versions of the PVST equivalent?
- Are any of the pictures in the PVST culturally-biased?
- What is the youngest age at which the PVST can be used?

ACKNOWLEDGEMENTS

Mark Toomer kindly recorded all the spoken sentences for the tests.

REFERENCES

- Anthony, L., van Hees, J., Nation, I. S.P. (Forthcoming). The vocabulary sizes of young native-speakers of English.
- Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Banerji, N., Gupta, V., Kilgarriff, A., & Tugwell, D. (2012). *Oxford Children's Corpus: a Corpus of Children's Writing, Reading, and Education*.
https://www.sketchengine.co.uk/documentation/raw.../Beebox_final.docx.
- Biemiller, A. (2010). *Words Worth Teaching: Closing the Vocabulary Gap*. Columbus: McGraw-Hill.
- Biemiller, A. (2005). Size and sequence in vocabulary development. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and Learning Vocabulary: Bringing Research into Practice* (pp. 223-242). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, 93(3), 498-520.
- Dale, E., & O'Rourke, J. (1976). *The Living Word Vocabulary*: Field Enterprises Educ. Corp., distributed exclusively by DOME, Inc., 1169 Logan Ave, Elgin 1160120.
- McKay, P. (2006). *Assessing Young Language Learners*. Cambridge: Cambridge University Press.
- Nation, I.S.P. & Anthony, L. (2016). Measuring vocabulary size. In *Handbook of Research in Second Language Teaching and Learning*, Volume III, E. Hinkel (Ed.) New York: Routledge.

- Parent, K. (2012). The most frequent English homonyms. *RELC Journal*, 43(1), 69-81.
- Read, J., & Chapelle, C. (2001). A framework for second language vocabulary assessment. *Language Testing*, 18(1), 3-32.
- Zhang, X. (2013). The 'I Don't Know' option in the Vocabulary Size Test. *TESOL Quarterly*, 47, 790-811.

NOTES

Comments/Suggestions/Bug Fixes

All new editions and bug fixes are listed in the revision history below. However, if you find a bug in the program, or have any suggestions for improving the program, please let me know and I will try to address the issues in a future version.

This software is available as 'freeware' according to the license below. It is important for my funding to hear about any successes that people have with the software. Therefore, if you find the software useful, please send me an e-mail briefly describing how it is being used.

CITING/REFERENCING THE PVST

Use the following method to cite/reference *PVST* according to the APA style guide:

Anthony, L. and Nation, I.S.P. (YEAR OF RELEASE). *Picture Vocabulary Size Test* (Version VERSION NUMBER) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>

For example if you download *PVST 1.0.0*, which was released in 2017, you would cite/reference it as follows:
Anthony, L. and Nation, I.S.P. (2017). *Picture Vocabulary Size Test* (Version 1.0.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>

Note that the APA instructions are not entirely clear about citing software, and it is debatable whether or not the "Available from ..." statement is needed. See here for more details:
<http://owl.english.purdue.edu/owl/resource/560/10/>

LICENSE FOR THE PICTURE VOCABULARY SIZE TEST (PVST)

Picture Vocabulary Size Test (PVST) 1.0 and any minor updates issued by AntLab Solutions (collectively 'the Software')

TERMS GOVERNING THE USE OF THE SOFTWARE

The Software is protected by copyright and must not be used, displayed, modified, adapted, distributed, transmitted, transferred or published or otherwise reproduced in any form by any means other than strictly in accordance with the terms set out below. By installing the Software, you agree to be bound by the terms of the license. This PVST License ("License") is made between AntLab Solutions, Tokyo, Japan as licensor, and you, as licensee, as of the date of your use of the Software. The Software is in use on a computer when it is loaded into the RAM or installed into the permanent memory of that computer, e.g., a hard disk or other storage device.

1. License Material

These terms govern your use of the Software but not including subsequent versions (e.g. PVST 2.0').

2. License Grant

AntLab Solutions grants to you a personal non-exclusive non-transferable license ('the License') to use the Software in the following specific contexts.

a) Non-Commercial (Freeware) Use:

You may use the software for non-profit purposes on more than one computer or on a network so long as you are the sole user of the Software. (A "network" is any combination of two or more computers that are electronically linked and capable of sharing the use of a single software program.) You are not permitted to sell, lease, distribute, transfer, sublicense, or otherwise dispose of the Software, in whole or in part, for any form of actual or potential commercial gain or consideration.

b) Commercial Evaluation (Trial) Use:

You may evaluate (trial) the software for commercial purposes for a period of no more than fourteen (14) days from the date of download on more than one computer or on a network so long as you are the sole user of the Software.

c) Commercial Use

When you pay the commercial license fee established by AntLab Solutions, you may use the software for non-profit or commercial purposes on more than one computer or on a network so long as you are the sole user of the Software. (A "network" is any combination of two or more computers that are electronically linked and capable of sharing the use of a single software program.) You will obtain a separate license for each additional user of the Software (whether or not such users are connected on a network). You are not permitted to sell, lease, distribute, transfer, sublicense, or otherwise dispose of the Software, in whole or in part, for any form of actual or potential commercial gain or consideration.

3. Termination

You may terminate this License at any time by uninstalling the Software and deleting it. The License will also terminate if you breach any of the terms of the License.

4. Proprietary Rights

The Software is licensed, not sold, to you. AntLab Solutions reserves all rights not expressly granted to you. Ownership of the Software and its associated proprietary rights, including but not limited to patent and patent applications, are retained by AntLab Solutions. The Software is protected by the copyright laws of Japan and by international treaties. Therefore, you must comply with such laws and treaties in your use of the Software. You agree not to remove any of AntLab Solutions' copyright, trademarks, and other proprietary notices from the Software.

5. Distribution

Except as may be expressly allowed in Section 2, or as otherwise agreed to in a written agreement signed by both you and AntLab Solutions, you will not distribute the Software, either in whole or in part, in any form or medium.

6. Transfer and Use Restrictions

You may not sell, license, sub-license, lend, lease, rent, share, assign, transmit, telecommunicate, export, distribute or otherwise transfer the Software to others, except as expressly permitted in this License Agreement or in another agreement with AntLab Solutions. You may not modify, reverse engineer, decompile, decrypt, extract, or otherwise disassemble the Software.

7. Warranties

ANTLAB SOLUTIONS MAKES NO WARRANTIES WHATSOEVER REGARDING THE SOFTWARE AND IN PARTICULAR, DOES NOT WARRANT THAT THE SOFTWARE WILL FUNCTION IN ACCORDANCE WITH THE ACCOMPANYING DOCUMENTATION IN EVERY COMBINATION OF HARDWARE PLATFORM OR SOFTWARE ENVIRONMENT OR CONFIGURATION, OR BE COMPATIBLE WITH EVERY COMPUTER SYSTEM. IF THE SOFTWARE IS DEFECTIVE FOR ANY REASON, YOU WILL ASSUME THE ENTIRE COST OF ALL NECESSARY REPAIRS OR REPLACEMENTS.

8. Disclaimer

ANTLAB SOLUTIONS DOES NOT WARRANT THAT THE SOFTWARE OR SERVICE IS FREE FROM BUGS, DEFECTS, ERRORS OR OMISSIONS. THE SOFTWARE OR SERVICE IS PROVIDED ON AN "AS IS" BASIS AND ANTLAB SOLUTIONS MAKES NO OTHER WARRANTIES OR CONDITIONS, EXPRESS OR IMPLIED, WITH RESPECT TO THE SOFTWARE INCLUDING WITHOUT LIMITATION THE IMPLIED WARRANTIES OR CONDITIONS OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

9. Limitation of Liability

ANTLAB SOLUTIONS WILL HAVE NO LIABILITY OR OBLIGATION FOR ANY DAMAGES OR REMEDIES, INCLUDING, WITHOUT LIMITATION, THE COST OF SUBSTITUTE GOODS, LOST DATA, LOST PROFITS, LOST REVENUES OR ANY OTHER DIRECT, INDIRECT, INCIDENTAL, SPECIAL, GENERAL, PUNITIVE OR CONSEQUENTIAL DAMAGES, ARISING OUT OF THIS LICENSE OR THE USE OR INABILITY TO USE THE SOFTWARE OR SERVICE. IN NO EVENT WILL ANTLAB SOLUTIONS'S TOTAL AGGREGATE LIABILITY (WHETHER IN CONTRACT (INCLUDING FUNDAMENTAL BREACH), WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY, INTELLECTUAL PROPERTY INFRINGEMENT OR OTHER LEGAL THEORY) WITH REGARD TO THE SOFTWARE AND/OR THIS LICENSE EXCEED THE LICENSE FEE PAID BY YOU TO ANTLAB SOLUTIONS. FURTHER, ANTLAB SOLUTIONS WILL NOT BE LIABLE FOR ANY DELAY OR FAILURE TO PERFORM ITS OBLIGATIONS UNDER THIS LICENSE AS A RESULT OF ANY CAUSES OR CONDITIONS BEYOND ANTLAB SOLUTIONS' REASONABLE CONTROL

10. Jurisdiction

These terms will be governed by Japanese law and the Japanese courts shall have jurisdiction.

KNOWN ISSUES

None at present

REVISION HISTORY

1.1.0

New features:

- 1) Test item sets are now not automatically included with the main test engine. Test item sets should be downloaded separately and placed in the "tests" folder. This feature reduces the size of the main program considerably.
- 2) Test item sets can now be loaded manually from the file menu.

1.0.0

This is the first version of the program

Copyright 2017 Laurence Anthony. All rights reserved.