SHOAIZ SHARAKHIMOV AND ULUGBEK NURMUKHAMEDOV
Uzbekistan and United States

Assessing Learners’ Productive Vocabulary Knowledge: Formats and Considerations

Vocabulary learning is an incremental process. Vocabulary knowledge, especially for second-language learners, may develop across a lifetime. For example, an advanced student in Uzbekistan told us that he recently learned that popular and famous have positive connotations, while notorious refers to being famous for something negative. An international graduate student in the United States complained that although she recognizes many words in English, especially when she reads books and journal articles, watches videos, or listens to audio materials, she cannot easily retrieve words effectively from her memory while speaking or writing. Because of this challenge, it takes her a lot of time to prepare for academic presentations or write academic papers for her graduate-level courses. Learning the meaning alone of new words does not guarantee word mastery—although it is an important first step.

Second-language learners need to know many features about a word because word knowledge is multifaceted and includes the following essential variables: form, meaning, and use. According to Nation (2013), form primarily involves recognizing a new word’s spelling and pronunciation. When learners see or hear a recently learned word, they should be able to recognize the word’s spelling or pronunciation and word parts (e.g., interest – interesting – interested). Meaning is about understanding a word’s synonyms, antonyms, associated words, and polysemy (having a large number of meanings). Use focuses on the ability to use a word in context by keeping in mind its lexicogrammatical nature and collocation (the frequent co-occurrence of two or more words).

In addition to form, meaning, and use, there exists in word knowledge a receptive and productive duality. An English language learner may recognize the meaning of a recently learned word in a reading passage (receptive) but might not be able to spell the word accurately, pronounce it intelligibly in speech, or use it correctly in an essay or in a conversation (productive). Unlike receptive vocabulary, productive vocabulary “involves knowing a lexical item [a word] well enough to produce it when it is needed to encode communicative content in speech or...
Assessing learners’ productive vocabulary involves unpredictable cases and/or inconsistent situations due to the multifaceted nature of word knowledge.

A learner using vocabulary—such as single words (arm, baby), compound words (armchair, babysit), phrasal verbs (sit through, sit down), collocations (cross your arms, sit tight), and idioms (cost [someone] an arm and a leg; sit on your hands)—in speech and/or writing needs to attend to the multiple aspects of vocabulary knowledge. To use vocabulary in productive ways, a learner should be able to do the following (in which T = talking and W = writing):

• pronounce a word intelligibly (T)
• spell the word correctly and write it legibly (W)
• identify the word’s derivation(s) or inflection(s), based on a context (T and W)
• consider the word’s grammatical function(s) to produce a syntactically accurate phrase or sentence (T and W)
• select a meaning sense and a semantic association to convey the intended meaning (T and W)
• apply collocational knowledge accurately to communicate the intended meaning (T and W)
• distinguish formal and informal words or phrases and be aware of when to use and not to use them (T and W)

This list does not encapsulate all aspects of productive vocabulary knowledge. Teachers with experience in providing feedback on their students’ vocabulary use in writing or speech might have noticed that it is sometimes difficult to pinpoint one aspect of word knowledge. The reason is that assessing learners’ productive vocabulary involves unpredictable cases and/or inconsistent situations due to the multifaceted nature of word knowledge. For example, if there are 15 students in a class, there could be 15 different and yet distinctive errors in relation to the use of the target word in speech or writing. Thus, identifying test formats that assess learners’ productive vocabulary use can be difficult. Some traditional test formats, such as multiple-choice and matching, are well suited for measuring learners’ receptive vocabulary skills (e.g., form–meaning connection, receptive vocabulary size) but are not the best option for measuring productive vocabulary skills. To help teachers track their students’ expressive vocabulary acquisition, this article presents five test formats that English language teachers can use or adapt to gauge their students’ productive vocabulary knowledge.

BEST PRACTICES FOR ASSESSMENT

Before we describe the test formats, we believe that teachers need to be aware of the following considerations. First, it is important to identify the purpose of the test. Teachers need to ask themselves which of the test formats they would like to use and for what purpose. It should be noted that some of these productive test formats assess learners’ vocabulary use through writing and others via speaking. Second, these test formats are not meant for use in achievement tests or in other forms of high-stakes assessment. Instead, they should be used for formative assessments that are not graded but provide learners an opportunity to review unit words in productive ways; although there is no formal grading, teachers have a chance to provide learners with feedback. Such formative assessments offer opportunities to create positive washback; washback refers to the
This article presents five test formats that English language teachers can use or adapt to gauge their students’ productive vocabulary knowledge.

influence “tests have on classroom practices—in particular, the effects they have on teaching and learning” (Wall 2012, 1). Third, when deciding on a test format, teachers need to take test practicality into consideration. In language assessment, practicality refers to the administrative issues involved in creating a test, including “costs, the amount of time it takes to construct and to administer, ease of scoring, and ease of interpreting/reporting the results” (Mousavi 2009, 518).

Finally, the purpose of a productive vocabulary test is to tap into a learner’s vocabulary use and determine how well the learner can use words in speaking and writing. To do so, teachers need to identify—as mentioned earlier—the type of productive vocabulary knowledge they aim to assess (e.g., collocations, meaning senses). Identifying the purpose will help them select an appropriate test format. This decision is also useful for test validity, which is the “extent to which a test measures what it is supposed to measure” (Bailey and Curtis 2015, 34). For example, if a student gets a high score in a vocabulary quiz at the end of a unit, we can infer that that score indicates good knowledge of the words that have been taught in the unit. Validity helps teachers understand how a construct—in this case productive vocabulary—is defined and what elements this construct entails. Thus, teachers should measure learners’ productive vocabulary knowledge by having them use words in speech or writing as opposed to giving them a multiple-choice or a matching test.

**FIVE TEST FORMATS FOR PRODUCTIVE VOCABULARY ASSESSMENT**

1. **Gap-Fill Test**

In gap-fill (or cloze) tests, students are presented sentences with missing words or sections and asked to fill in the blanks with correct words. For example:

1. COVID-19 is an infectious _______ that has quickly spread worldwide, causing a pandemic.

Note that several answers (word options), such as disease and virus, may be written as a correct response in Sentence 1. Although this flexibility may be pleasing to teachers, it is not ideal in a test because having more than one answer complicates scoring; test items with only one plausible answer are preferred. In a classroom context, teachers may want their students to fill in the gap by using a word they have learned in a unit/chapter. To elicit words that were previously taught in class, teachers should present sentences that have a limited context so that there is only one correct answer. To minimize answer options, learners also have to know the number of gaps they need to answer. In the following examples, learners need to supply one word in Sentence 2, two words in Sentence 3, and three words in Sentence 4:

2. I read an interesting book on how to cope _______ difficult people.

3. She is not going to _______ _______ with their smoking any longer.

4. I do not want to go outside now. Look! It’s raining _______ _______ _______.

[answer key: (2) with; (3) put up; (4) cats and dogs]

**Scoring.** To describe the scoring method, we decided to use Sentence 1, “COVID-19 is an infectious _______ that has quickly spread worldwide, causing a pandemic.” As
The purpose of a productive vocabulary test is to tap into a learner’s vocabulary use and determine how well the learner can use words in speaking and writing.

previously discussed, plausible answers for this item include disease and virus. To simplify our explanation, let us assume that a teacher taught the word disease in a previous unit, and the students, when quizzed, were expected to write disease in the sentence. Methods for scoring gap-fill tests should be discussed in advance with students. For example, a correct answer could receive 2 points, a partially correct answer 1 point, and an incorrect answer 0 points. Partially correct responses can include misspelling and grammatical inaccuracy:

• Misspelling: The learner misspells the target word (e.g., disese or dizease, both incorrect); such a spelling mistake should be pointed out to the student. If the word is not recognizable (e.g., diziz), then the student has not learned the spelling of the word, though the same student may know the pronunciation of the target word.

• Grammatical inaccuracy: A learner writes a correct word but in a wrong grammatical form (e.g., diseased or diseases, neither of which fits grammatically in the given sentence).

After students receive their test scores, teachers should encourage them to go back to the unit and review the vocabulary items.

Benefits and Considerations
Overall, a gap-fill is a valid test format for assessing productive vocabulary because learners are asked to write appropriate words in blanks. Additionally, gap-fill tests are simple to create, though sentences should be chosen carefully. To create context-specific items, teachers can extract example sentences from their course textbooks, which may be familiar to learners, or from freely available online learner dictionaries that are easy to understand (see a list of learner dictionaries in Nurmukhamedov 2012).

2. Productive Vocabulary Levels Test

The productive version of the Vocabulary Levels Test (VLT) is recommended for teachers who wish to learn how many words learners recognize in each frequency level—that is, how many words are known from the 1,000, 2,000, and 3,000 frequency levels. Nation (2013) provides a detailed overview of popular frequency lists, including the VLT. Although the VLT was originally designed to measure learners’ receptive vocabulary size, Laufer and Nation (1999) modified the test format and introduced the productive version of the test, calling it the VLT Productive.

In the VLT Productive, learners are presented with a sentence in which one word is purposely left incomplete; only the first two or three letters of the incomplete word are provided. Learners are asked to read the sentence and supply the rest of the letters of the incomplete word. For example, Sentence 5 below prompts a learner to write ease (to complete the word disease).

5. COVID-19 is an infectious dis________ that has quickly spread worldwide, causing a pandemic.

6. COVID-19 is an infectious vi________ that has quickly spread worldwide, causing a pandemic.

7. I read an interesting book on how to co________ with difficult people.

8. The program’s main objective is to imp________ educational standards.
This is a controlled productive vocabulary format because while learners take the test, they are not just producing random words; instead, they are prompted to supply letters that generate a word that is semantically appropriate in a meaningful sentence context.

Scoring. Scoring the items may range from 0 (wrong) to 1 (correct). Spelling errors may be penalized, depending on the teacher’s course policy.

Benefits and Considerations
To assess how well target words have been learned, teachers can use this format and integrate the previously learned words into their own test. As with other test formats, the sentences in this test can be adapted from a course textbook or learner dictionaries.

While creating the blanks, teachers can use the following procedure: the first two letters may be shown for words of up to five letters (e.g., co for cope; vi for virus). For words that have more than five letters, the first three letters may be shown (e.g., dis for disease; imp for improve).

3. Vocabulary Knowledge Scale

With the Vocabulary Knowledge Scale, learners are asked to demonstrate how well they know a given set of words. This test format, originally proposed by Wesche and Paribakht (1996), has been used in many research projects that examine English language learners’ depth of vocabulary knowledge. The following modified version of the test was created by Folse (2006).

In Table 1, the word that is being tested is toil, and the learner is required to demonstrate knowledge of the word by filling in the blanks.

There are three parts in each test item. If the learner indicates unfamiliarity with the target word by selecting #1, the learner does not need to complete #2 or #3. If the learner knows the meaning of the word, then he or she defines the target word in #2, either in English with a synonym or (if the teacher allows it) in his or her native language with a translation. After completing #2, the learner answers #3, which requires writing a sentence using the target word. In sum, the test items indicate how well a learner knows the target word.

Scoring. Learners can earn up to 2 points per test item:

- 0 points: The learner indicates that he or she does not know what the word means (#1).
- 1 point: A correct meaning is demonstrated in #2 (e.g., toil means “work hard” or “work with great effort”).
- 1 point: A correct sentence is generated using the target word in #3 (e.g., “Farm workers have to toil in the field all day.”).

1. I don’t know what the word toil means.

2. I know this word. It means ________________________________.
   [Provide an English synonym or a translation in your native language.]

3. I can use this word in a good example sentence. Here is my example sentence:
   ____________________________________________________.
   [If you do #2, you must also do #3.]

Table 1. Modified Vocabulary Knowledge Scale (adapted from Folse 2006, 281)
Teachers can formulate their own way of scoring to make the scores informative and meaningful for their own context.

**Benefits and Considerations**

This test format is primarily used to assess learners’ productive vocabulary via writing. Using this test format, teachers can also elicit vocabulary knowledge by interviewing learners. Interviewing tests oral production of the target items rather than written production; however, it should be noted that interviewing each student in class—depending on the class size—may be time-consuming and may even be impossible to achieve with large classes.

The Vocabulary Knowledge Scale comes in several modified formats. Interested teachers can explore the application of a different version of the test in a book titled *Focus on Vocabulary* (Schmitt, Schmitt, and Mann 2011), which introduces academic words via reading and writing activities and then uses a modified Vocabulary Knowledge Scale as a pretest and posttest to assess learners’ vocabulary gains before and after instruction.

**4. Meaningful Sentences**

In this test format, learners are presented with target words and asked to orally produce meaningful utterances using those words. This test format can be used in conversation classes where listening and speaking are the primary focus. During the test, teachers can either play a prerecorded audio prompt or read out loud the instructions shown in Table 2 (as an illustration of the test directions, only two words are listed).

During the ten-second pause, learners may take notes on a piece of paper or mentally prepare their sentence. During the 20-second pause, learners are asked to record their sentences using either their cell phone or a voice-recorder device.

Learners should be told explicitly that the goal is to convincingly inform the teacher—through the production of learner-generated meaningful sentences—that they have learned the target words. For the word *dictionary*, for example, a meaningful sentence may include a full definition or be a complete sentence. Instead of saying, “A dictionary is a book” (which is not a thorough definition), learners might produce sentences like the following:

- Student 1: “A dictionary gives the meaning of words.”
- Student 2: “A book that have many word.”
- Student 3: “A dictionary has useful information about words. I can check word definitions and example sentences in a dictionary.”
- Student 4: “We have a lot of dictionaries in the classroom, and we use them a lot while reading a new text.”

---

**Instructions:** Turn on your recorder. Do not turn off your recorder until you are told to do so. You will hear five words. Each word will be read twice. After hearing the word twice, you will have ten seconds to mentally prepare a meaningful sentence and then 20 seconds to record your sentence before the next word is read. Your sentence may contain a different form of the word provided. The test will begin now.

The first word is *dictionary*. **Dictionary.** You have ten seconds to prepare a meaningful sentence [ten-second pause]. You now have 20 seconds to state your sentence [20-second pause].

The second word is *gadget*. **Gadget.** You have ten seconds to prepare a meaningful sentence [ten-second pause]. You now have 20 seconds to state your sentence [20-second pause].

---

**Table 2. Test instructions for Meaningful Sentences**
The student utterances for the target word dictionary indicate that all four students know what it means and can—to some extent—define it in a meaningful way. Students 1 and 2 produced simple sentences. Student 2 was able to communicate the meaning in simpler terms and with some grammar errors, but these did not impede comprehension. All four example sentences vary in terms of content, target vocabulary use, and grammatical accuracy.

**Scoring.** Teachers can assess the quality of student-generated sentences by using the simple scorecard shown in Table 3 (adapted from McFeely and Nurmukhamedov 2014).

This test scorecard contains the target words, cells for a binary score (0 points for an unsuccessful sentence; 1 point for a meaningful sentence), and space for teacher comments. Scoring criteria for student-generated sentences may include variables such as

- the acceptability of the given meaning (coherent or incoherent),
- grammaticality (accurate or inaccurate),
- pronunciation (intelligible or unintelligible).

The scorecard allows teachers to score the learner-generated utterances and provide feedback that pinpoints specific errors made by learners. Because the target words are meant to be derived from the course unit or lesson, learners may go back to the unit and learn more about the words upon receiving the teacher’s comments.

**Benefits and Considerations**

We offer two caveats regarding this test format. After students record their sentences using teacher-provided recorders or their own digital devices, transferring students’ files to a teacher’s computer may be time-consuming, especially in large classes or for teachers with limited computer skills. To improve the practicality of the test, a teacher needs to find an efficient way to make this transfer, such as having learners email their audio files to the teacher.

The second caveat is that learners might hear one another speaking while recording their sentences. Learners should be seated far enough apart that they do not hear one another’s sentences, which can provide an unintended opportunity for cheating or disturb students while taking the test.

### 5. Context-Dependent Collocation Test

This test format measures whether a learner can produce collocations—two- or three-word combinations of words that frequently occur together. Collocations pose challenges for learners because “learners often know words but are unable to use them effectively because they do not know their collocates” (Webb and Sasao 2013, 269–270).

For example, strong and powerful are frequently used words and synonymous with each other; however, powerful collocates with car and

<table>
<thead>
<tr>
<th>#</th>
<th>Word</th>
<th>Score</th>
<th>Teacher Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dictionary</td>
<td>0/1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>gadget</td>
<td>0/1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>component</td>
<td>0/1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>direct</td>
<td>0/1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>essential</td>
<td>0/1</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3. Scoring a Meaningful Sentences task (adapted from McFeely and Nurmukhamedov 2014)*
engine, while strong collocates with coffee and views (see McCarthy, O’Keefe, and Walsh 2010, for more information about this topic).

In this test format, students are asked to read a short passage that contains two or three awkward or wrong collocations (i.e., miscollocations). After learners have read the passage, they are asked to revise the miscollocations. The following passage contains one verb–noun miscolligation and two adjective–noun miscolligations. In this task, learners need to replace the verb (solve) and adjectives (fantastic and nice) with appropriate collocates—but without changing other parts of the sentences.

Not all people have the chance to study in another country. When students in Kuwait finish high school, they need to (1) solve a decision: What should I do next? Most of them decide to go abroad to study because they believe that study abroad could have a (2) fantastic role in their future career. In addition, most parents believe that study abroad can have a (3) nice impact on their children.

In the example passage, awkward collocations have already been underlined, and wrong collocates have been bolded to help learners focus on the words that need to be replaced. For more-advanced students, teachers may refrain from underlining and bolding.

After students read the passage and identify correct collocations, they write their answers in the Revised Collocation column on the collocation worksheet (see Table 4). Learners may write as many possible collocations as they want to in the Revised Collocation column, but they should be reminded to keep the context in mind while revising the awkward collocations.

**Scoring.** Correct collocations receive a score of 1, and wrong collocations receive 0. Students who write additional correct collocations may receive extra points. A teacher should keep in mind that a student-generated collocation may be correct out of context, but it must be correct in the context of the test passage. For example, the collocation notorious singer in isolation may sound correct, but in a specific context would be incorrect:

- **Incorrect:** “Everybody respected her because she was a notorious singer.”
- **Correct:** “Everybody respected her because she was a famous singer.” (This usage is appropriate because the sentence context has a positive connotation.)

To help students understand the role of context in producing collocations, teachers can go over students’ responses with the whole class.

**Benefits and Considerations**

Writing teachers may find this test format valid and useful because it encourages learners to read a short passage and produce correct collocations based on the ideas presented in the passage. The test is easy to create and administer. Teachers may adapt passages from textbooks or write original texts based on the types of collocations taught in the unit/

<table>
<thead>
<tr>
<th>#</th>
<th>Awkward Collocation</th>
<th>Revised Collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>solve a decision</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>fantastic role</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>nice impact</td>
<td></td>
</tr>
</tbody>
</table>

[Possible answers: (1) make a decision; (2) crucial / important / vital role; (3) significant / major / big impact]

Table 4. Collocation worksheet
Use these test formats in formative assessments (ungraded exercises) as opposed to high-stakes assessments (achievement tests).

chapter. Passage length can vary, depending on the number of collocations teachers would like to insert in a passage. Some miscollocations can be derived from students’ writing (and/or from their speech) if teachers notice awkward collocation use by students.

FOOD FOR THOUGHT

Before implementing the proposed test formats, teachers should set a goal: Which productive vocabulary skill do they plan to measure? This is important because different test formats elicit different aspects of productive word knowledge (e.g., spelling, collocation knowledge, usage of the words in speaking or writing). Target words taken from a course lesson/unit are beneficial; recently learned words are of greatest value to learners’ vocabulary development; and unit-/lesson-based words tend to appear in high-stakes summative examinations, such as midterm and achievement tests. The test formats provided in this article can be used to diagnose learners’ productive vocabulary knowledge before new words in a unit are introduced. Furthermore, these test formats are classroom friendly and promote formative assessment. Additionally, contextual information is necessary to understand or produce the target word. Read and Chapelle (2001) argue that context-dependent tests inform teachers whether a learner knows a word. We suggest that teachers consider the following recommendations when using the test formats in their respective contexts:

• Check your students’ proficiency and vocabulary size. We believe that the suggested test formats may be used for students with intermediate language proficiency levels and above. We suggest that colleagues who apply Common European Framework of Reference for Languages (CEFR) levels in their contexts use these tasks with learners in the B1 through C2 levels. The reason is that the learners in these levels recognize the most frequent 2,500 and 5,000 words in English (see Moore 2020). Because the suggested test formats are designed to elicit responses through writing or speaking, learners need to have enough vocabulary to communicate their meaning. Eager teachers can always experiment with A2 level learners and find out which of the suggested formats may be suitable for their students.

• Brush up on word knowledge. Explain to your students that recognizing 2,500 words does not equate to being able to use those words in speech or writing. They need to be aware of essential factors that account for word knowledge. Review some of the key terms, such as word parts, collocations, and grammatical functions, together with your students. For example, word parts for interest include interests, interesting, interested, uninteresting, and uninterested. The word interest typically collocates with adjectives such as strong, keen, and growing, and with verbs such as express, show, take, and stimulate.

• Train for the test format(s) and review tests with students. Before implementing one of the test formats, carefully introduce the test format and its purpose. Provide several examples and a few practice examples for students to try on their own or with peers before testing. After students have taken a productive vocabulary test, encourage them to review their test. Allow them to discuss each other’s responses in the test and possibly provide each other with peer feedback. Encourage students to go back to the course textbook and review target words.
• **Use the test formats for formative assessment.**
  Most important, use these test formats in formative assessments (ungraded exercises) as opposed to high-stakes assessments (achievement tests). The goals are to (a) improve learners’ productive vocabulary knowledge by providing them with multiple opportunities to use the words in speaking and writing and (b) raise their awareness that receptive and productive vocabulary knowledge require different sets of skills and word knowledge.

**CONCLUSION**

The five test formats in this article can be a useful addition to your vocabulary assessment toolbox. Talk with colleagues about how they assess English language learners’ productive vocabulary and what works for them. As we mentioned earlier, choosing a test format needs to be done in light of your learning objectives and purpose. In addition, some tests are more practical than others in regard to time for administration and scoring. Through trial and error and some reflection, language teachers will learn which test formats work best in their respective contexts and for their specific students.

**REFERENCES**


**Shoaziz Sharakhimov** teaches at Tashkent Medical Academy in Uzbekistan. His research interests include second language assessment and vocabulary.

**Ulugbek Nurmukhamedov** teaches in the MATESL program at Northeastern Illinois University in the United States. His research interests include second language vocabulary and computer-assisted vocabulary learning.