

# Development of an Online Adaptive Vocabulary Test System

Wen-shuenn Wu

Department of Foreign Languages & Literature

Chung Hua University

Hsinchu, Taiwan

[wswu@chu.edu.tw](mailto:wswu@chu.edu.tw)

**Abstract:** The prototype of an *online adaptive vocabulary test system* (OAVTS) will be demonstrated. The system, the first stage of a web-based adaptive English placement test platform, is now being developed to help universities group students into classes which match their English proficiency. In a computer-adaptive language test (CALT), each examinee takes a unique test that is tailored to his ability level. It is generally believed that CALT is faster, more reliable and more cost-effective. OAVTS is expected to help teachers or test makers quickly prepare and edit their vocabulary test items online. Some practical issues such as the classification of the word list, definitions of item codes, selection of the test-stopping rule, and design of test makers' interface are also discussed.

## Introduction

It has been a tendency that more and more universities in Taiwan would like to give freshmen an English placement test so that they can be grouped into appropriate levels of English competence. In all likelihood, a web-based English placement test is the only solution to the laborious and time-consuming task of classifying freshmen in the first week of the fall semester. Since placements tests are generally not strictly tied to a particular text, computer-adaptive language test (CALT) development has received great attention in the field of placement test. Compared to a traditional paper-and-pencil test or even a web-based, linear test, CALT is faster, more precise and effective in terms of placing students into different proficiency level. With the help of corpus-based word list, it is more statistically objective for test makers to decide the difficulty level of vocabulary test items. Therefore, a prototype of *Online Adaptive Vocabulary Test System* (OAVTS), the first stage of a web-based adaptive English placement test platform, was firstly developed. The long-term goal of this project is to set up a web-based adaptive English placement test platform in which hundreds of university students can take an English placement test simultaneously. Test results, shown immediately after students take it, can be easily stored and retrieved for further analyzed and studied.

## Basic Design of the Online Adaptive Vocabulary Test System

This OAVTS is divided into three interfaces – student interface for examinees to take an English placement test, teacher interface for teachers or test makers to prepare, edit and upload their test items, and administrator interface for system administrators to manage all the test items. The following CALT aspects were considered when OAVTS was first developed.

### Difficulty Classification of English Words

The following dictionaries were referred to determine the difficulty level of English words.

- A. Collins COBUILD English Dictionary for Advanced Learners (2001) (abbrev. as *Collins*)
- B. Macmillan English Dictionary for Advanced Learners of American English (2002) (abbrev. as *Macmillan*)
- C. Longman Dictionary of Contemporary English (2003) (abbrev. as *Longman*)

The lexical database of OAVTS was classified into six difficulty level.

- Level 0: includes basic and the most frequently used words, such as *the, a, of, you, this, is, are*, etc.
- Level 1: includes words in the frequency band 5 and 4 of *Collins*.
- Level 2: *Longman* labels the most frequently used spoken words (S1~S3) and written words (W1~W3). If they have not been categorized into Level 1, then they belong to Level 2.
- Level 3: are words of the frequency band 3 of *Collins* if they do not belong to *Longman's* S1~S3 and W1~W3, plus *Macmillan's* two-star entries and three-star entries if they are not in Level 1 and Level 2.
- Level 4: are *Macmillan's* one-star entries plus *Collins* frequency band 2.
- Level 5: are words labeled as frequency band 1 of *Collins*.

The breakdown of the six-level English words into stems and derivatives is shown in Table 1.

level	0	1	2	3	4	5
stems	141 words	1,567 words	1,548 words	1,179 words	2,670 words	6,198 words
derivatives(word family)		3,864 words	3,435 words	2,583 words	6,038 words	13,066 words

**Table 1:** Number of words in each difficulty level of OAVTS

### Definitions of Item Codes

Clearly defined item codes of web-based adaptive English placement test can easily serve a double function

of both placement and diagnosis. Item codes, consisting of 15 numbers or alphabets, were used in OAVTS to carefully state the features and purposes of each test item. The representation of each code is defined as follows:

- 1st code: type (vocabulary, grammar, listening, reading, etc.)
- 2nd ~ 4th codes: difficulty parameter (these three codes are reserved for the item difficulty of item response theory (IRT) logistic model)
- 5th code: specification (identification, comprehension, collocation, spelling, contextual clue, synonymous clue, etc.)
- 6th ~ 7th codes: category (general, business, science, medicine, law, environment, etc.)
- 8th code: part of speech (noun, verb, adjective, adverb, phrasal verb, etc.)
- 9th code: invariable N means “number”.
- 10th ~ 15th codes: serial numbers

### Selection of the Test-Stopping Rule

The following two factors were mainly considered when the test-stopping rule of OAVTS was decided. First, since most vocabulary test formats are multiple-choice, it is highly possible that examinees may guess the answer. Second, if examinees stay in a certain difficulty level of OAVTS for quite a while, it may mean that they are not good enough to be upgraded into a higher level, nor are they bad enough to be downgraded to a lower level. Figure 1 illustrates the test-stopping rule of OAVTS.

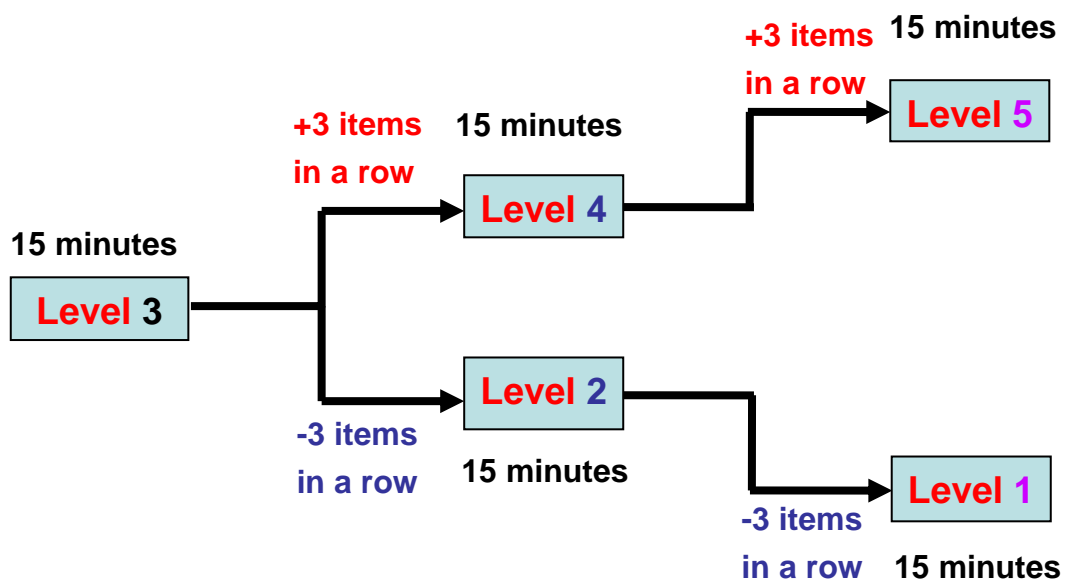


Figure 1: The stopping-rule of OAVTS

All examinees are given a test item from Level 3. If they answer three items correctly in a row, they will be given more difficult test items from Level 4. Similarly, if examinees missed three questions in a row, they will be given easier questions from Level 2. In addition, timer will be set once examinees take the test. If they stay in a certain level for 15 minutes and get stuck there, the test will stop automatically. Since OAVTS was first created for English placement test, only time spent in each level and probability of lucky guess were considered. However, stopping rules of CALT are more complicated than we imagined.

While the convention of each examinee’s first item is usually elicited from medium-difficulty level of item bank, the selection of the test-stopping rule can be various in CALT. According to Wainer (1990, p.114), computer-adaptive test (CAT) terminates when one or more of the following stopping rules is met:

- A. when a target measurement precision level has been achieved
- B. when a pre-selected number of items has been given
- C. when predetermined amount of time has elapsed

Dunkel (1999) also reminded test developers to put type of test (e.g. fixed-length or variable-length) and stopping rules in the documentation in CAT. In the future, if IRT-based logistical models are used in OAVTS, a well-thought test-stopping rule will be needed to achieve a satisfactory level of precision.

### Design of Test Makers’ Interface



Figure 2: A screen shot of test makers’ interface of OAVTS

A network-based, user-friendly interface for test makers was designed so that test makers can concentrate on preparing test items. Test makers do not have to memorize the details of the item codes; they only have to choose a proper one from pull-down menus. A foolproof device was also put in OAVTS. Test developers have to assign a difficulty level to a test item before they can submit the question to the test bank.

## **Test Security**

When a browser-based adaptive test is administered, the security of student interface is quite important. Students' computers will be configured to show full screen so that they can not use any of those pull-down menus or toolbars. Both the right-click feature of a mouse and backspace function of the keyboard will also be disabled so that examinees can not go back to previous test items. Finally, the URL of test items will be only valid for a limited period of time to prevent examinees from typing those URLs after the exam to make any changes.

## **Limitation of the Study**

The prototype of OAVTS was created specifically to give English teachers in my university to have a vocabulary test-making platform on which a corpus-based word list was built. Therefore, in the beginning IRT statistical framework was not taken into consideration. It is definitely the first limitation of this project. However, IRT-based CALT is an interdisciplinary field of study, involving at least three different areas of knowledge – English, computer science, and mathematics. Language experts from the department of English are responsible for providing question items, professionals of computer science handle programming, user interface, test security, and server maintenance, and specialists in mathematics deal with IRT logistical models and statistical analysis. It really needs a team work. Second, even before IRT theory was applied to OAVTS, the difficulty level of each word is not the only factor deciding the difficulty of test items. Sentence patterns, length, grammar complexity and test makers' previous knowledge toward test questions all affect the difficulty level of test items.

## **Conclusion**

CALT has become a growing trend of standardized foreign language assessment. Online adaptive vocabulary test is the first step toward a web-based adaptive English placement test. As Larson (1999) states, computer-adaptive testing has great potential to provide testing research data. With the implementation of OAVTS, plus a carefully designed English proficiency survey and a well-defined item codes, test developers can easily collect and analyze test results to make a more reliable and valid CALT and English teachers can also explore test

results to find out implications for pedagogical improvement.

## References

- Dunkel, P. (1999). Considerations in developing or using second/foreign language proficiency computer-adaptive tests. *Language Learning & Technology*, 2(2), 77-93. Retrieved August 7, 2003, from <http://lt.msu.edu/vol2num2/article4/index.html>
- Larson, J. (1999). Considerations for testing reading proficiency via computer-adaptive testing. In M. Chalhoub-Deville (Ed.), *Issues in computer-adaptive testing of reading proficiency* (pp. 71-90). New York: Cambridge University Press.
- Wainer, H., Dorans, N., Flaugher, R. Green, B., Mislevy, R., Steinberg, L. & Thissen, D. (1990). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Lawrence Erlbaum Associates.

## Acknowledgments

The research in this article was funded by a faculty research grant (CHU-92-H-004, November 1, 2003 – October 31, 2004) from Chung Hua University, Hsinchu, Taiwan.