

Analyzing the Appropriateness of Vocabulary in O-NETs of Grade 12 Thai Students

การวิเคราะห์ความเหมาะสมของคำศัพท์ในข้อสอบ O-NET ของนักเรียนไทย ชั้นมัธยมศึกษาปีที่ 6

Supika Nirattisai*
ศุภิกา นีรัตติสัย*

General Education, Phuket Rajabhat University
ศึกษาทั่วไป มหาวิทยาลัยราชภัฏภูเก็ต

ABSTRACT

The objectives of this present research were to examine the appropriateness of words used in the English O-NETs of grade 12 Thai students and to explore the consistency of the word difficulties used in these tests. The instruments used in this research were the 2008-2010 English O-NETs. Check Words Up Software was used to classify words based on their difficulty levels, 1st-20th word levels. In order to answer the research questions, basic statistics were applied. The results revealed that the words used in each English O-NET were quite suitable to grade 12 Thai students, based on the expected level of their English proficiency set by the Basic Education Core Curriculum (A.D. 2008). However, there were some inappropriate words in some parts of each test. In terms of the consistency, 2008-2010 O-NETs seemed to use equal levels of word difficulties.

ARTICLE INFO

Article history:

Received 6 July 2016

Received in revised form

7 November 2016

Accepted 16 November 2016

Available online

25 December 2017

Keywords:

O-NET

Appropriate (เหมาะสม)

Vocabulary (คำศัพท์)

English proficiency

(ความสามารถภาษาอังกฤษ)

Grade 12 Thai students

(นักเรียนไทยชั้นมัธยมศึกษา

ปีที่ 6)

* Corresponding author

E-mail address: supika_mew@hotmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบความเหมาะสมของคำศัพท์ภาษาอังกฤษและความสม่ำเสมอของระดับความยากของคำศัพท์ภาษาอังกฤษที่ใช้ในข้อสอบ O-NET ภาษาอังกฤษสำหรับนักเรียนไทยในระดับชั้นมัธยมศึกษาปีที่ 6 เครื่องมือที่ใช้ในงานวิจัยนี้ คือ ข้อสอบ O-NET ภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 6 ในปี 2551-2553 และโปรแกรม Check Words Up ได้นำมาใช้แบ่งแยกคำศัพท์ตามระดับความยาก “จากระดับ 1 ถึง ระดับ 20” สถิติเบื้องต้นทางคณิตศาสตร์ได้นำมาใช้เพื่อตอบคำถามวิจัยข้างต้น ผลการวิจัย พบว่า คำศัพท์ที่ใช้ในข้อสอบ O-NET ภาษาอังกฤษ แต่ละปี (2551-2553) มีระดับค่อนข้างเหมาะสม โดยอ้างอิงจากระดับความสามารถทางภาษาอังกฤษของนักเรียนไทยในระดับชั้นมัธยมศึกษาปีที่ 6 ที่ระบุในหลักสูตรแกนกลางปี 2551 อย่างไรก็ตาม คำศัพท์ภาษาอังกฤษบางคำในข้อสอบ O-NET แต่ละปี ยังคงพบว่ามีความไม่เหมาะสมกับระดับความสามารถทางภาษาอังกฤษของนักเรียนในระดับชั้นดังกล่าว ความสม่ำเสมอของระดับความยากคำศัพท์ในข้อสอบ O-NET ภาษาอังกฤษแต่ละปี พบว่า มีความสม่ำเสมอในระดับค่อนข้างดี

Introduction

Ordinary National Education Tests (O-NETs) are developed by the National Institute of Education Testing Service responsible for developing tools of educational evaluation (NIETS, 2013, p.1). The main aims of the tests are: 1) to assess the achievement of grade 6, 9, and 12 Thai students in the eight core subjects – Thai Language, Mathematics, Science, Social Science, Religion & Culture, Foreign Languages, Health & Physical Education, Art, and Career & Technology, 2) to give teachers information of students’ English language proficiency in order to enhance their teaching styles, and 3) to measure the quality of Thai education (NIETS, 2013). “Among O-NETs in the 8 different fields” English O-NET is one of the important achievement test. This is because English language is a medium of international communication in various fields, such as businesses, education, industry (Crystal, 1997); having a high level of English proficiency would allow Thai students to communicate with speakers of other languages.

The English O-NETs of grade 12 Thai students was the focus of this present study, because of the significance of the test scores for grade 12 students; English O-NET scores constitute part of Thai national university entrance examination.

English O-NETs are conducted based on the Basic Education Core Curriculum (The Ministry of Education Thailand, 2008, p. 1). In order to evaluate the appropriateness of English O-NETs of grade 12 Thai students, it was crucial to know their 4 expected English qualities set by the Ministry of Education Thailand. These 4 mentioned qualities were; 1) Language for Communication, 2) Language and Culture, 3) Language and the Relationship with other learning areas, and 4) Language and Relationship with Community and the World (The Ministry of Education Thailand, 2008, p. 252-281). Of all 4, the first quality “*Language for Communication*” was the quality researched in this research paper.

The expected English knowledge of grade 12 Thai students based on the first type “*Language for Communication*” was as follows. Grade 12 Thai students should be able to comprehend the various types of written and spoken texts including feature articles, entertainment articles, news, announcements, advertisements, poems, and skits; when they read or listen to them and they should be able to express their opinions with appropriate reasons; they should be able to appropriately speak and write to exchange or give information about themselves, daily-life situations, news issues, and other interest around them.

Among many effective ways to assess the quality of tests, in this present study, vocabulary analysis was chosen to be the method for considering the test quality. There has been a lot of research that has revealed vocabulary as being a good indicator for measuring learners’ language proficiency in all main skills: reading, listening, writing, and speaking (e.g., Laufer, 1998; Nation & Meara, 2002). Learners with higher English proficiency would have a larger vocabulary level. Thus, vocabulary used in English tests for advance learners needs to be more difficult than vocabulary for beginners and intermediate learners. In other words, a good quality test should contain appropriate vocabulary, in terms of levels of word difficulty, based on learners’ English proficiency level.

In order to construct a high quality English O-NET, an achievement test, for grade 12 Thai students, vocabulary used in this test should be one of the main qualities being carefully considered. Obviously, vocabulary is embedded in all parts of the English O-NET. Although, this test has only a small amount of items in the vocabulary section, the other sections such as conversation, reading and writing also contain a large number of words. Therefore, using too many higher difficulty words than their expected English knowledge based on the Basic Education Core Curriculum in grade 12 English O-NET would affect the assessment of their English achievement.

This research aimed to evaluate the appropriateness of vocabulary difficulty used in English O-NETs. The results could reveal whether words applied in the tests were suitable to grade 12 Thai students’ expected English level.

What amount of vocabulary is sufficient for grade 12 Thai students?

The core curriculum indicates that grade 12 Thai students are expected to acquire vocabulary of about 3,600 - 3,750 word families in different word levels. A word family is a group of words under one head word; one word family is comprised of the base word (head word), derived forms, and reflected forms which we can recognize without learning each word members individually (Nation, 1993). The words such as accept, accepts, accepted, accepting, acceptable, acceptably, acceptability, unacceptable, unacceptably are considered as one word family with the head word “accept”. According to the amount of words presented in the core curriculum, it tends to be insufficient for grade 12 Thai

students. The explanation of this claim was as follows.

How many word families should be the ideal number for grade 12 Thai students based on the Basic Education Core Curriculum? Hu & Nation (2000) and Schmitt, Jiang & Grabe (2011)'s studies showed that to be able to comprehend written or spoken texts without any assistance, to know at least 98% of running words in texts (98% word coverage) – 2 unknown words appearing on every 100 words. This number might sound too many for language learners. However, the research conducted by Schmitt, Jiang, and Grabe (2011) revealed that with knowing 98% of words in a text, learners would obtain only 68% comprehension of text. Thus, learners who have more than 2% of the unknown words would struggle with a full understanding of texts. The 98% word coverage was the target number for this present research.

To consider the suitable vocabulary amount of grade 12 Thai students, the ideal word coverage of 98% together with the expected level of their English knowledge announced by the Thai Education Institute was considered. With 98% coverage as the adequate figure, to be able to read graded reader series produced by various publishers, learners would need to know vocabulary generally starting at 200 word families up to 3,800 word families or at around 3,000 word families plus proper nouns (Schmitt, N. & Schmitt, D, 2014). Therefore, grade 12 students need to acquire words more than 3,000 word families because they are expected to read texts which are more difficult than just graded readers.

Based on Nation (2006), the adequate number of words for learners to read English authentic texts including novels and news was 8,000 - 9,000 word families (plus proper nouns) to comprehend these kinds of texts. However, it took smaller amount of vocabulary at 5,000–6,000 word families for learners to have a full understanding of spoken texts including everyday conversation and unscripted spoken English. In order to comprehend the text when learners watch children's movies, they tended to need more words, at about 6,000–7,000 word families to cope with them; however, the ability to comprehend children's movies are not the expected English knowledge of grade 12 students revealed in the core curriculum.

As can be seen, the target number of words has to be higher than the word amount given in the Basic Education Core Curriculum which is about three thousand five hundred word families. In order to reach the expected English level based on the Basic Education Core Curriculum, grade 12 Thai students need a higher level of vocabulary knowledge. In conclusion, the ideal vocabulary used in this present study were 5,000 and 8,000 word families. Five thousand word families would be the lowest vocabulary level for grade 12 Thai students to comprehend the texts in the dialogue section of

English O-NETs and eight thousand word families would be the smallest number to fully understand the texts in the other parts of English O-NETs, such as reading, vocabulary and the writing section.

Research Questions

1. Are the words used in the English O-NETs appropriate in terms of spoken texts and written texts for assessing grade 12 Thai students' English skills?
2. Is there a consistency of word difficulties across the three English O-NETs (2008-2010)?

Research Methodology

Data Collection

This study included three papers of English O-NETs for grade 12 Thai students from the years 2008-2010. These papers were developed by the Thai National Institute of Educational Testing Service, and were aimed to measure Thai education' quality in English, as well as to assess Thai students' English achievement levels for improving classroom techniques used in schools (NIETS, 2013). As previously stated, there were only three O-NETs (2008-2010) used in this study. The reasons were; 1) there are not any O-NETs after 2010 available, the responsible organization would keep them for some time before publishing them, and 2) even though 2005-2007 O-NETs are available, they are not based on the same Basic Education Core Curriculum (A.D. 2008).

The 2008 O-NET consisted of four main parts: conversation, error identification, writing, and reading. For 2009 and 2010 O-NETs, they were comprised of three main parts which were conversation, writing, and reading.

Data Analysis

1. Data analysis tool

Check Up Words Software was created for the present research by Pillai and Nirattisai (2016). It was used to analyze words in any kind of written and spoken texts into 20 different levels, based on the frequency levels occurring in the BNC (the British National Corpus). The 1st 1000 word level consists of the words with the highest frequency level occurring in texts; it implies that the words in this level are the easiest group among all 20 levels. The 20th 1000 word level is the highest infrequency one. This word group includes the most difficult words of all. The words in levels higher than the 20th 1000 level are very rarely found in common texts, except texts in specific fields of studies.

The specific characteristics of this software are as following. Firstly, according to the reliability test, this software has 100% correctness of analyzing words in each word level, except the unsuitable typing words and spaces. Secondly, this software spends quite a short time in analyzing each of 100 words. Finally, the data produced by this software presented in terms of both a total number and a percentage number of words in the 20 word levels. Moreover, the results from the software are convenient for researchers to have a closer look at the levels of each word because they are highlighted in different colors based on the word levels.

2. Data obtaining processes

The O-NET files were imported into the software, after which the software would run and export the analyzed data. This data presented the total numbers and percentages of words occurring in each word level. Furthermore, the software also revealed the bands in different colors to facilitate the data analysis, by allowing the researcher to easily see the levels of each word difficulty in the whole texts. All data analyzed by the software would be used for further analysis.

3. Analysis Processes

To answer the first research question, summary statistics were used to compute the summation of words used in: 1) *the 1st–5th 1,000 word levels + the 4 kinds of words (proper nouns, compound words, marginal words and abbreviations)* – the knowledge of at least 5,000 word families was considered to be a minimum sufficient level for grade 12 Thai students to comprehend spoken texts based on their expected English level presented in the Basic Education Core Curriculum, and 2) *the 1st–8th 1,000 word levels + the 4 kinds of words* – the lexical knowledge of 8,000 head words was a minimum adequate level for them to read all types of written texts based on grade 12 students' expected English proficiency showed in the Basic Education Core Curriculum.

To answer the second research question, the differences between the smallest and the largest percentage of each word level in the three O-NETs (2008-2010) were calculated. After that, the data from the earlier calculation was used to answer whether the gaps occurring between the three O-NETs are large or not.

Results and Discussion

Research Question 1: *Are the words used in the English O-NETs appropriate in terms of spoken texts and written texts for assessing grade 12 Thai students' English skills?*

Words used in each O-NET were divided into 20 word levels based on their difficulties: 1st–20th 1000 word levels. The first 1st 1000 word level was the highest frequency level words used in spoken and written texts; the 20th 1000 word level are the lowest frequency words used in spoken and written texts. In other words, the 1st 1000 word level contains the easiest words among the 20 word levels and the 20th word level consists of the most difficult words among them.

Here are some examples of words in some levels.

| | |
|---|--|
| 0 – 1000: big, the, use | 10001 – 11000: exonerate, papal, singe |
| 2001 – 3000: alcohol, Asia, wild | 12001 – 13000: biomass, panama, squirm |
| 4001 – 5000: drought, obey, stump | 14001 – 15000: camphor, ingratiante, nectarine |
| 6001 – 7000: ascend, innovative, wand | 16001 – 17000: belabour, liverwort, protozoan |
| 8001 – 9000: exhaustion, muscular, trivia | 18001 – 19000: hubcap, preterm, stingray |

As mentioned in the introduction above, grade 12 Thai students need at least a word knowledge of between 1st–5th 1000 word levels to comprehend texts in the conversation part of O-NETs. The word knowledge of 1st–8th 1000 word levels was the lowest one for them to comprehend the texts in other parts of O-NETs such as reading, vocabulary, and writing. Therefore, these numbers were used as target criteria to decide whether the words used in English O-NETs were appropriate for assessing their English abilities or not.

Table 1 shows the levels of word difficulties used in the 2008 O-NET. This test consisted of five parts: conversation, error identification, writing, vocabulary, and reading passages.

Table 1. Words used in the 2008 O-NET

| | Number of words in the test (%) | | | | | |
|--|---------------------------------|----------------------|---------|------------|---------|-----------|
| | Conversation | Error Identification | Writing | Vocabulary | Reading | All parts |
| The 4 kinds of words* | 3.40 | 0.64 | 1.90 | 3.53 | 2.67 | 2.66 |
| 1 st 1000 word family | 89.55 | 82.58 | 76.43 | 69.02 | 77.08 | 80.35 |
| 2 nd 1000 word family | 3.64 | 7.74 | 12.65 | 15.29 | 10.22 | 8.91 |
| 3 rd 1000 word family | 1.82 | 5.16 | 3.29 | 4.71 | 4.32 | 3.56 |
| 4 th 1000 word family | 0.10 | 0.97 | 1.56 | 2.35 | 2.22 | 1.44 |
| 5 th 1000 word family | 0.29 | 0.97 | 0.00 | 1.57 | 0.64 | 0.53 |
| 1 st – 5 th levels (+ the 4 kinds of words) | 98.80 | - | - | - | - | - |
| 6 th 1000 word family | 0.29 | 0.65 | 1.04 | 2.35 | 0.51 | 0.67 |
| 7 th 1000 word family | 0.38 | 0.00 | 0.17 | 0.39 | 0.64 | 0.43 |
| 8 th 1000 word family | 0.48 | 0.00 | 0.69 | 0.00 | 0.32 | 0.37 |
| 1 st – 8 th levels (+ the 4 kinds of words) | - | 98.71 | 97.73 | 99.21 | 98.62 | 98.92 |
| 9 th 1000 word family | 0.10 | 0.65 | 1.04 | 0.00 | 0.13 | 0.29 |
| 10 th 1000 word family | 0.00 | 0.00 | 0.17 | 0.39 | 0.13 | 0.11 |
| 11 th 1000 word family | 0.00 | 0.32 | 0.17 | 0.00 | 0.57 | 0.29 |
| 12 th 1000 word family | 0.00 | 0.00 | 0.52 | 0.00 | 0.13 | 0.13 |
| 13 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 | 0.05 |
| 14 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 15 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 16 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 17 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.25 | 0.11 |
| 18 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.39 | 0.06 | 0.05 |
| 19 th 1000 word family | 0.00 | 0.00 | 0.17 | 0.00 | 0.00 | 0.03 |
| 20 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Upper levels | 0.00 | 0.33 | 0.18 | 0.00 | 0.00 | 0.06 |
| 9 th – upper levels | 0.10 | 1.30 | 2.25 | 0.78 | 1.40 | 1.12 |
| Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

* The 4 kinds of words included proper nouns (e.g., Malee, Bangkok, KFC), transparent compounds (e.g., *football*, *skincare*, *teacup*), marginal words (e.g., *Um*, *Er*, *Gosh*), and abbreviations.

The results revealed that 98.92% of all words in the 2008 O-NET (plus the four kinds of words: proper nouns, compound words, marginal words, and abbreviations) were in the 1st-8th 1000 word levels, a minimum knowledge for learners to comprehend all types of reading texts, and only 1.12% of words were in the higher levels.

For each part of the test, 98.80% of words in the conversation part consisted of the words in the 1st-5th 1000 word levels, the target word knowledge for learners to truly understand spoken texts. In error identification, writing, vocabulary, and reading part, 98.71%, 97.73%, 99.21%, and 98.62% of words in those parts, respectively, were in the 1st-8th 1000 word families, the essential amount required for grade 12 Thai students to comprehend the types of written texts.

In total, most words used in the overall and separate parts of the 2008 O-NET were in the appropriate word levels for grade 12 Thai students. It had only a small number of words above their expected English knowledge.

It was interesting to note that 2.25% of words in the writing section, 1.30% in the error identification part, 0.78% in the vocabulary, and 1.40% in reading were in the higher word levels than the target one at 8th 1000 word level. Point one percent of words in the conversation part was above the target level at 5th 1000 word level.

According to the number above, the writing part had the highest percentage at 2.25% when compared with other parts. It meant that there were 2.25 unsuitable words occurring in every 100 words. In other words, grade 12 Thai students would know about 97.75% of all words in this writing part based on their expected English proficiency; it was below 98% of text coverage which was considered to be an adequate amount for learners to comprehend each text (Hu & Nation, 2000). Thus, the high number of 2.25% inappropriate words, could affect their comprehension. Thus, some inappropriate words in this part needed to be revised in order to make a better achievement test. Words above grade 12 Thai students' expected English level could appear in this O-NET, but the number of them should be carefully considered.

Table 2 presents the levels of words used in the 2009 O-NET. This test was separated into four parts which were conversation, writing, vocabulary, and reading passages section. An error identification section disappeared from this O-NET; it became to be a part of the writing section.

Table 2. Words used in the 2009 O-NET

| | Number of words in the test (%) | | | | |
|--|---------------------------------|---------|------------|---------|-----------|
| | Conversation | Writing | Vocabulary | Reading | All parts |
| The 4 kinds of words* | 6.65 | 2.94 | 0.89 | 3.84 | 3.97 |
| 1 st 1000 word family | 86.19 | 86.12 | 82.67 | 75.06 | 82.42 |
| 2 nd 1000 word family | 4.98 | 7.03 | 11.56 | 7.94 | 7.10 |
| 3 rd 1000 word family | 0.83 | 0.92 | 1.78 | 4.91 | 2.22 |
| 4 th 1000 word family | 0.31 | 1.28 | 3.11 | 2.19 | 1.45 |
| 5 th 1000 word family | 0.21 | 0.55 | 0.00 | 1.34 | 1.01 |
| 1 st – 5 th levels (+ the 4 kinds of words) | 99.17 | - | - | - | - |
| 6 th 1000 word family | 0.31 | 0.37 | 0.00 | 0.91 | 0.51 |
| 7 th 1000 word family | 0.00 | 0.43 | 0.00 | 1.45 | 0.31 |
| 8 th 1000 word family | 0.52 | 0.31 | 0.00 | 0.53 | 0.41 |
| 1 st – 8 th levels (+ the 4 kinds of words) | - | 99.95 | 100.00 | 98.17 | 99.65 |
| 9 th 1000 word family | 0.00 | 0.06 | 0.00 | 0.15 | 0.07 |
| 10 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 11 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.23 | 0.07 |
| 12 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.08 | 0.02 |
| 13 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.15 | 0.05 |
| 14 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.23 | 0.07 |
| 15 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 16 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 17 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.08 | 0.24 |
| 18 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 19 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.15 | 0.05 |
| 20 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Upper levels | 0.00 | 0.00 | 0.00 | 0.77 | 0.25 |
| 9 th – upper levels | 0.00 | 0.06 | 0.00 | 1.84 | 0.82 |
| Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

* The 4 kinds of words include proper nouns (e.g., Malee, Bangkok, KFC), transparent compounds (e.g., *football*, *skincare*, *teacup*), marginal words (e.g., *Um*, *Er*, *Gosh*), and abbreviations.

In Table 2, 99.65% of words in the 2009 O-NET were in the 1st-8th 1000 word levels; almost 100% of them were comprised of the appropriate vocabulary for grade 12 Thai students. In terms of each separated part, 99.17% of words in the conversation part were in the 1st-5th 1000 word levels. 99.95%, 100.00%, and 98.17% of words in writing, vocabulary, and reading part, respectively, were in the 1st-8th 1000 word levels. It could be said that the whole test and individual parts contained quite reasonable word difficulties for grade 12 Thai students. The number of appropriate words used in the test was very close to the ideal number at 100%.

In the reading part, there was the highest percentage of words (1.84%) above the 8th 1000 word level. However, this number did not tend to have much effect on their reading comprehension. Based on this number, grade 12 Thai students would understand 98.17% of words in the reading part; it was higher than 98% text coverage considered to be the adequate number for comprehending texts (Hu & Nation, 2000).

There was another interesting result shown in this O-NET. In the vocabulary part, there were no words above 4th 1000 word levels. However, the objective of this part was to test grade 12 Thai students' vocabulary knowledge. The words used in this part should have a variety of word difficulties. To make a better quality test, more difficult words should be added to this part.

Table 3 illustrates the word levels used in the 2010 O-NET. This test contained four parts: conversation, writing, vocabulary, and reading part.

Table 3. Words used in the 2010 O-NET

| | Number of words in the test (%) | | | | |
|--|---------------------------------|---------|------------|---------|-----------|
| | Conversation | Writing | Vocabulary | Reading | All parts |
| The 4 kinds of words * | 7.97 | 1.75 | 0.55 | 0.19 | 2.38 |
| 1 st 1000 word family | 86.84 | 85.32 | 69.40 | 74.89 | 81.01 |
| 2 nd 1000 word family | 3.14 | 9.35 | 14.75 | 11.57 | 9.16 |
| 3 rd 1000 word family | 1.09 | 1.75 | 3.83 | 5.46 | 3.09 |
| 4 th 1000 word family | 0.60 | 0.97 | 2.73 | 2.10 | 1.40 |
| 5 th 1000 word family | 0.12 | 0.78 | 1.09 | 2.70 | 1.38 |
| 1 st – 5 th levels (+ the 4 kinds of words) | 99.76 | - | - | - | - |
| 6 th 1000 word family | 0.00 | 0.06 | 1.09 | 0.20 | 0.15 |
| 7 th 1000 word family | 0.00 | 0.00 | 4.92 | 0.39 | 0.37 |
| 8 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.33 | 0.12 |
| 1 st – 8 th levels (+ the 4 kinds of words) | - | 100.00 | 98.36 | 97.83 | 99.06 |
| 9 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 10 th 1000 word family | 0.24 | 0.00 | 0.00 | 0.59 | 0.27 |
| 11 th 1000 word family | 0.00 | 0.00 | 0.55 | 0.13 | 0.07 |
| 12 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.20 | 0.07 |
| 13 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 14 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.33 | 0.12 |
| 15 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 16 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.53 | 0.20 |
| 17 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.26 | 0.10 |
| 18 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 19 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 20 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Upper levels | 0.00 | 0.00 | 1.09 | 0.14 | 0.10 |
| 9 th – upper levels | 0.24 | 0.00 | 1.64 | 2.18 | 0.93 |
| Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

* The 4 kinds of words include proper nouns (e.g., Malee, Bangkok, KFC), transparent compounds (e.g., *football*, *skincare*, *teacup*), marginal words (e.g., *Um*, *Er*, *Gosh*), and abbreviations.

According to Table 3, 99.06% of words in 2009 O-NET were in the 1st-8th 1000 word levels; a small number (0.94%) was above the expected word levels. In the conversation part, 99.76% of words were in the 1st-5th 1000 word levels. In the writing, vocabulary, and reading part, 100.00%, 98.36%, and 97.83% of words, respectively, were in the 1st-8th 1000 word levels. Overall, the words in this test mostly reached the ideal number at 100%. In other words, the words in the test were quite appropriate for grade 12 Thai students, except the words used in the reading part.

In the reading part, 2.18% of words above the 8th 1000 word level. It meant that grade 12 Thai students would understand about 97.32% of text coverage; it was lower than the adequate number at 98% coverage (Hu & Nation, 2000). This would affect their comprehension. However, among the inappropriate words used in this part, it was found that some of them were repeated many times. For example, the word “blanching” in the 16th 1000 word level occurred in this part five times and the word “asparagus” in the 14th 1000 level also occurred five times. Thus, the high number revealed in this part might be caused by the repeating words.

Moreover, according to the words used in the item 67 of the reading part, it revealed that the word in the first option was in the 12th 1000 word level (homicide), but the target word was in the 11th 1000 word level (manslaughter). As mentioned earlier, the words in a higher word level tend to have a higher difficulty than the one in a lower word level. Thus, it could be seen that the word used in the first option (it was a correct choice of this item), was more difficult than the target word. Thus, the difficulty levels of a target word and words in choices should be carefully considered. This is the example of item 67.

67. The word “**manslaughter**” (line 10) can be replaced by ____.

- | | |
|---------------------|-----------------|
| 1. homicide | 2. imprisonment |
| 3. irresponsibility | 4. Disability |

In summary, the words used in the whole test were fine. However, the words in the reading part still needed some revisions.

Research Question 2: *Is there a consistency of word difficulties across the three English O-NETs (2008-2010)?*

To see the consistency of word difficulties across the three English O-NETs (2008-2010), the relationship between the percentage of text coverage and the percentage of text comprehension should be explained. According to Schmitt, Jiang & Grabe's study (2011), learners who know the words between 95% – 98% (plus proper nouns) of written texts obtain comprehension levels at 60% - 68%. It meant that the decrease of 1% text coverage will drop learners' comprehension of between 0.63% - 0.69% (formulas: $60/95 = 0.63$; $68/98 = 0.69$); the average number would be about 0.66. Thus, a high inconsistency level of word difficulties across tests would represent a high unstandardized level of O-NETs.

To analyze consistency of word difficulties across the three O-NETs, the range between the highest and lowest percentage of words used in the three O-NETs should be very small. The smaller the ranges between the three O-NETs were, the higher the consistency of the word difficulties in them was.

In order to easily analyze the consistency of the O-NETs, 20 word levels were categorized into three groups: high-, middle-, and low- frequency words, based on the study of Norbert Schmitt & Diane Schmitt (2014). The high-frequency words are those which appeared most frequently in everyday language (1st - 3rd 1000 word levels). The low-frequency words are words rarely found in texts (9th word level and upper) and the middle-frequency words are words which we do not always see in texts, however we rarely find (4th – 8th 1000 word levels).

Table 4 presents the comparison of word difficulties used in the three English O-NETs (2008-2010).

Table 4. Comparing the word difficulties in 2008-2010 English O-NETs.

| | Number of words (%) | | | Max | - Min | Range |
|-----------------------------------|---------------------|------------|------------|-------|---------|-------|
| | 2008 O-NET | 2009 O-NET | 2010 O-NET | | | |
| The 4 kinds of words * | 2.66 | 3.97 | 2.38 | 3.97 | - 2.38 | 1.59 |
| 1 st 1000 word family | 80.35 | 82.42 | 81.01 | 82.42 | - 80.35 | 2.07 |
| 2 nd 1000 word family | 8.91 | 7.10 | 9.16 | 9.16 | - 7.10 | 2.06 |
| 3 rd 1000 word family | 3.56 | 2.22 | 3.09 | 3.56 | - 2.22 | 1.34 |
| High frequency words | 92.82 | 91.74 | 93.26 | 93.26 | - 91.74 | 1.52 |
| 4 th 1000 word family | 1.44 | 1.45 | 1.40 | 1.45 | - 1.40 | 0.05 |
| 5 th 1000 word family | 0.53 | 1.01 | 1.38 | 1.38 | - 0.53 | 0.85 |
| 6 th 1000 word family | 0.67 | 0.51 | 0.15 | 0.67 | - 0.15 | 0.52 |
| 7 th 1000 word family | 0.43 | 0.31 | 0.37 | 0.43 | - 0.31 | 0.12 |
| 8 th 1000 word family | 0.37 | 0.41 | 0.12 | 0.41 | - 0.12 | 0.29 |
| Middle frequency words | 3.44 | 3.69 | 3.42 | 3.69 | - 3.42 | 0.27 |
| 9 th 1000 word family | 0.29 | 0.07 | 0.00 | 0.29 | - 0.00 | 0.29 |
| 10 th 1000 word family | 0.11 | 0.00 | 0.27 | 0.27 | - 0.00 | 0.27 |
| 11 th 1000 word family | 0.29 | 0.07 | 0.07 | 0.29 | - 0.07 | 0.22 |
| 12 th 1000 word family | 0.13 | 0.02 | 0.07 | 0.13 | - 0.02 | 0.09 |
| 13 th 1000 word family | 0.05 | 0.05 | 0.00 | 0.05 | - 0.00 | 0.05 |
| 14 th 1000 word family | 0.00 | 0.07 | 0.12 | 0.12 | - 0.00 | 0.12 |
| 15 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | - 0.00 | 0.00 |
| 16 th 1000 word family | 0.00 | 0.00 | 0.20 | 0.20 | - 0.00 | 0.20 |
| 17 th 1000 word family | 0.11 | 0.24 | 0.10 | 0.24 | - 0.10 | 0.14 |
| 18 th 1000 word family | 0.05 | 0.00 | 0.00 | 0.05 | - 0.00 | 0.05 |
| 19 th 1000 word family | 0.03 | 0.05 | 0.00 | 0.05 | - 0.00 | 0.05 |
| 20 th 1000 word family | 0.00 | 0.00 | 0.00 | 0.00 | - 0.00 | 0.00 |
| Upper levels | 0.06 | 0.25 | 0.10 | 0.25 | - 0.06 | 0.19 |
| Low frequency words | 1.12 | 0.82 | 0.93 | 1.12 | - 0.82 | 0.49 |

* The 4 kinds of words include proper nouns (e.g., Malee, Bangkok, KFC), transparent compounds (e.g., *football*, *skincare*, *teacup*), marginal words (e.g., *Um*, *Er*, *Gosh*), and abbreviations.

According to Table 4, in the group of the high frequency words (the 1st – 3rd 1000 word levels), the 2010 O-NET had the highest percentage of words (93.26%) among all three tests, followed by the 2008 O-NET (92.82%), and the 2009 O-NET (91.74%), respectively. The range between the highest and the lowest percentages of this word group was 1.52%.

The words in the middle frequency levels, the 4th–8th 1000 word levels, were used most in the 2009 O-NET (3.69%), followed by the 2008 O-NET (3.44%), and the 2010 O-NET (3.42%), respectively. The difference between the highest and the lowest percentages of this word group was 0.27%.

In the low frequency word levels (the 9th – upper levels), the 2008 O-NET contained the highest amount (1.12%); the second and last place were the 2010 O-NET (0.93%) and the 2009 O-NET (0.82%). The range between the highest and lowest percentage of this word group was 0.49%.

In total, the ranges between the highest and lowest percentage of the words in the high-, middle-, and low- frequency word levels were 1.52%, 0.27%, and 0.49%, respectively. Based on the information mentioned above, the decrease of 1% of known words would lower learners' reading comprehension by about 0.66%. Thus, the ranges happening in the high-, middle-, and low- frequency word group would slightly affect their comprehension at 1.00%, 0.18%, and 0.32%, respectively. In other words, the difficulty levels of words in 2008-2010 O-NETs were quite consistent.

Even though, there was the highest gap (1.52%) happening in the high frequency word group, this gap didn't seem to affect much on grade 12 Thai students' reading comprehension because the words in the 1st – 3rd 1000 word levels were the everyday words and grade 12 Thai students should be really familiar with most of those words.

The consideration should be more focused on the ranges occurring in the groups of middle- and low- frequency words. As stated in Schmitt and Diane Schmitt's study (2014), the intermediate and advanced learners tended to acquire vocabulary in middle- and low-frequency levels more than beginners and elementary learners. Thus, big ranges happening in these two frequency word levels could really affect learners' reading comprehension. According to the results in Table 4, there were only small spaces occurring in the middle- and low-frequency levels (0.27%, and 0.49%, respectively); it would affect grade 12 Thai students' reading comprehension of about 0.18% and 0.32%, respectively. This range seemed to be acceptable because it only slightly impacted their reading comprehension.

Conclusion

The purposes of this research were to explore the appropriateness of words used in the grade 12 English O-NETs and to reveal the consistency of words used across the English O-NETs. The results were summarized as follows.

1. In terms of the 2008 English O-NET, 98.92% of words were appropriate, based on grade 12 Thai students' expected English level. 98.80% of words in the conversation part, 98.71% in the error identification part, 97.73% in the writing part, 99.21% in the vocabulary part, and 98.62% in the reading part used the suitable word levels. Overall, most words in this O-NET were appropriate to the levels being tested. However, there were a high number of inappropriate words in the writing part.

2. For the 2009 English O-NET, 99.65% of words were suitable to grade 12 students. 99.17%, 99.95%, 100.00%, and 97.17% of words in the conversation, writing, vocabulary, and reading part, respectively, were considered appropriate. Obviously, there was an appropriateness of vocabulary used in the overall 2009 O-NET, except for the reading and vocabulary part. There were a number of rather difficult words used in the reading part, but there were much easy vocabulary in the vocabulary part.

3. For the 2010 English O-NET, There were 99.06% of words in the suitable word levels. 99.76% of words in conversation, 100.00% in writing part, 98.36% in vocabulary part, and 97.83% in reading part were revealed to be suitable for them. Apparently, the words used in this O-NET were appropriate for the students. Yet, the proportion of rather difficult words used in the reading part was relatively high. Furthermore, the word used in the first option of item 67, the correct choice, was more difficult than the target word.

4. Across the 2008 – 2010 O-NETs, the ranges between the highest and lowest percentage of the high-, middle-, low- frequency level were 1.52, 0.27, and 0.49, respectively. The results revealed the close ranges in the three frequency word groups. It meant that the difficulty levels of words in the three O-NETs were quite consistent.

Pedagogical and testing implication

This study showed that most of the words used in the English O-NETs for grade 12 Thai students were in the appropriate word levels, in terms of their difficulty, except a small number of words in some test sections. Besides, the difficulties of words in the tests (the 2008-2010 English O-NETs) are quite consistent.

The appropriateness and consistency of each English O-NETs needs to accurately reflect the expected level being tested.

Further Studies

This present research aimed to evaluate the appropriateness of vocabulary used in grade 12 English O-NETs in order to reveal the validity of this achievement test based on the Basic Education Core Curriculum (A.D. 2008). For further research, other effective methods should be applied to recheck the appropriateness of English O-NETs. Analyzing the correlations between students' O-NETs scores and vocabulary knowledge would be another instrument for considering the O-NETs' quality; according to Nation (1990) and Schmitt (2010), learners' English proficiency is related to their vocabulary knowledge.

References

- Crystal, D. (1997). **English as a global language**. Cambridge University Press: Cambridge.
- Hu, M. and Nation, P. (2000). Vocabulary density and reading comprehension. **Reading in a Foreign Language** 23(1): 403-430.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? **Applied Linguistic** 19(2): 255-271.
- Nation, I.S.P. (1990). **Teaching and learning vocabulary**. Boston, MA: Heinle and Heinle.
- Nation, P. (1993). Word Families. **International Journal of Lexicography** 6(4): 254-279.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening? **Canadian Modern Language Review** 63(1): 59-82.
- Nation, P., & Meara, P. (2002). Vocabulary. In N. Schmitt (Ed.), **An introduction to applied linguistics**. New York: Oxford University Press Inc.
- Schmitt, N. (2010). **Researching vocabulary: A vocabulary research manual**. Basingstoke: Palgrave Macmillan.
- Schmitt, N., Jiang, X. & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. **The Modern Language Journal** 95(1): 26-43.
- Schmitt, N. and Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. **Language Teaching** 47(4): 484-503.
- The Ministry of Education Thailand. (2008). **Learning Area of Foreign Languages** [Online]. Available: <http://www.skn.ac.th/kan2551.htm>
- The National Institute of Educational Testing Service (NIETS). (2013). **Ordinary National Educational Test** [Online]. Available: <http://www.niets.or.th>