

國立政治大學英國語文學系碩士在職專班碩士論文

指導教授：尤雪瑛博士

Advisor: Dr. Hsueh-ying Yu

2010 年至 2017 年英語科國中基測與會考克漏字測驗分析

Analysis of the Cloze Test Items in English Basic Competence Test and  
Comprehensive Assessment Program from 2010 to 2017

研究生：彭凡家 撰

Name: Fan-chia Peng

中華民國一百零七年六月

June, 2018

Analysis of the Cloze Test Items in English Basic Competence Test and  
Comprehensive Assessment Program from 2010 to 2017

A Master Thesis  
Presented To  
Department of English  
National Chengchi University



In Partial Fulfillment  
of the Requirements for the Degree of  
Master of Arts

By  
Fan-chia Peng  
June, 2018

## Acknowledgements

First of all, I would like to show my deepest gratitude to my thesis advisor, Dr. Hsueh-ying Yu of the Department of English of the National Chengchi University. With her inspiring and patient guidance, I could complete my thesis. Besides, when I was in the time of frustration, I could always have some relief from her words of wisdom and her consistently encouragement.

My sincere thanks also extend to my thesis committee members, Dr. Chieh-yue Yeh and Dr. Hsi-nan Yeh. With their insightful suggestions and comments on the manuscript, I could make my thesis more complete and persuasive.

I would like to show my appreciation to all my classmates in ETMA. I was grateful with their company. Thanks to them, I had more confidence to fight for the huge stress of the studies during the times of several exhausted but beautiful summers in ETMA. My special thank goes to Grace Hsu, the coder of the thesis, for her assistance to my research.

Lastly, I am much obliged to my husband Tsung-hsin Liu and my family for their support and love from the beginning to the end of this wonderful journey.

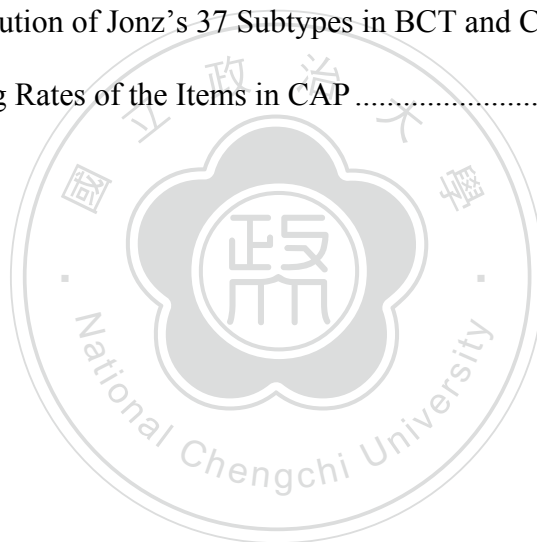
## TABLE OF CONTENTS

Acknowledgements.....	iii
Table of Contents .....	iv
List of Tables .....	v
List of Figures .....	vi
Chinese Abstract .....	vii
English Abstract .....	viii
<b>Chapter One: Introduction .....</b>	<b>1</b>
Background and Motivation .....	1
Significance of the Study .....	4
<b>Chapter Two: Literature Review .....</b>	<b>5</b>
Testing and Assessment in Curriculum .....	5
Approaches to Language Testing.....	6
The Development of Cloze Test .....	8
Measurement of the Readability.....	8
Measurement of the Reading Comprehension .....	9
Types of Modified Cloze Tests.....	10
What Kind of Language Skill Does Cloze Measure? .....	15
Relevant Research on Cloze Test in Taiwan .....	17
BCT and CAP .....	18
BCT .....	19
CAP .....	19
<b>Chapter Three: Methodology .....</b>	<b>21</b>
Materials .....	21

Instrument .....	22
Modification of Jonz's Subcategories .....	27
Data Analysis .....	35
<b>Chapter Four: Results and Discussion.....</b>	<b>39</b>
Results of the Five Categories .....	39
The Results of Jonz's 37 Subtypes in BCT and CAP .....	44
The Distribution of Jonz's 5 Categories in Each Year of BCT and CAP .....	49
Results of the Distribution of Jonz's 37 Subtypes in BCT and CAP .....	53
<b>Chapter Five: Conclusion.....</b>	<b>59</b>
Summary of the Major Findings .....	59
Implications.....	61
Limitations and Suggestions .....	62
References.....	63
Appendix A .....	67
Appendix B .....	68
Appendix C .....	71
Appendix D .....	76
Appendix E .....	77

## LIST OF TABLES

Table 3. 1 Distribution of Reading Passages and Test Items in Cloze Test .....	22
Table 3. 2 Jonz's (1990) Subtypes of Cloze Items .....	24
Table 3. 3 Modified Taxonomy of the Current Study .....	34
Table 4. 1 The Distribution of Items in the Five Categories .....	40
Table 4. 2 The Distribution of Jonz's Subtypes in BCT and CAP .....	45
Table 4. 3 The Distribution of Jonz's Five categories from 2010 to 2017 .....	50
Table 4. 4 The Numbers of Words and Readability in BCT .....	52
Table 4. 5 The Numbers of Words and Readability in CAP .....	52
Table 4. 6 The Distribution of Jonz's 37 Subtypes in BCT and CAP .....	54
Table 4. 7 The Passing Rates of the Items in CAP .....	56



## LIST OF FIGURES

Figure 2. 1 The Flowchart of Jonz's Interpretation of Bachman's Categories.....	13
Figure 3. 1 Data Analysis Procedure of the Study.....	35



國立政治大學英國語文學系碩士在職專班

碩士論文提要

論文名稱：2010 年至 2017 年英語科國中基測與會考克漏字測驗分析

指導教授：尤雪瑛博士

研究生：彭凡家

論文提要內容：

對台灣國中生而言，2001 年實施至 2013 年的國中基本學力測驗以及 2014 年開始實施的國中教育會考，是兩項重要的高中入學考試。其中，英文閱讀測驗題組中每年會有二篇的克漏字測驗，克漏字測驗能夠測驗受試者英文學習的整合能力，且在台灣被廣泛地使用於各項考試中，本研究主要目的在於：(1)研究從 2010 年至 2017 年基測與會考的克漏字試題主要測驗了哪些語言能力；(2)分析這 8 年的克漏字試題分別在基測與會考中測驗的能力為何，以及討論其相同與相異之處。本研究使用 Jonz (1990) 的克漏字分類法為其研究工具以探討其分類結果。

本研究結果顯示：(1) 2010 年至 2017 年基測與會考著重於受試者要能找出跨越句子與句子間文本線索的能力；(2) 2010 至 2013 年的基測，著重於檢視受試者時態句法的能力；(3) 2014 至 2017 年的會考，著重於檢視受試者篇章結構的能力。根據本研究發現，筆者針對未來研究方向及教學實務提出了建議。

關鍵字：克漏字測驗、國中教育會考、國中基本學力測驗



## Abstract

To every ninth-grade student of junior high school, the Basic Competence Test (BCT) from 1999 to 2013 and the Comprehensive Assessment Program (CAP) from 2014 are the two most important senior high school entrance examinations. Among the reading tests of both exams, two of them are the cloze tests. The cloze test targets at measuring the candidates' performances on the integrated ability in language, and it is widely used in the different testing contexts in Taiwan. The research thus aims to investigate: (1) from 2010 to 2017, what kind of language knowledge in the cloze tests were measured in BCT and CAP, and (2) the similarities and differences of the cloze item types tested and designed in BCT and CAP. To achieve the purpose, this research analyzed the cloze items in BCT and CAP in the past 8 years, and Jonz's taxonomy was utilized as the instrument to classify the target test items.

The analysis results of the present study are shown as follows: (1) the primary language knowledge tested from 2010 to 2017 was about the higher-order knowledge that focused on the ability to integrate the contextual clues across sentential level; (2) the most frequently-tested items in BCT was the syntactic knowledge in tense; and (3) the most frequently-measured item in CAP was the textual relationship in discourse. Based on the findings of this research, the researcher yielded suggestions for the future research and implications for teaching.

Key words: cloze test, Basic Competence Test, Comprehensive Assessment Program



## CHAPTER ONE

### INTRODUCTION

#### Background and Motivation

The Basic Competence Test for Junior High School Students (BCT, hereafter) was the main standardized entrance exam for every ninth-grade graduate to take from 1999 to 2013. In 2014, the Comprehensive Assessment Program for Junior High School Students (CAP, hereafter) replaced BCT as the new test type. The test result is of vital importance for it decides the future school to each graduate. Both BCT and CAP are the large-scale high-stakes national tests that require a large amount time of preparation for both test designers and test-takers. The test items have therefore received much attention.

To well understand BCT and CAP, it is essential to realize the purpose of both national exams and the principles of test construction. According to the Research Center for Psychological and Educational Testing at National Taiwan Normal University (RCPET, hereafter), the test construction of BCT and CAP are based on the Grade 1-9 Curriculum Guidelines which was announced in 2008. As for the purpose of both exams, they are summative achievement tests which aim to evaluate learners' academic acquisition after three-year junior high school education.

The relationship between official policy and the national summative assessment is highly related to EFL teaching in Taiwan. The 12-year basic education policy have implemented since 2014, and it was the year BCT was replaced by CAP. There is a change of the difficulty level in CAP. According to RCPET (Appendix A), BCT is designed at the level of easy, and CAP is at the level of appropriate which means the proportion of the difficulty is equally distributed among the test. The shift of difficulty from BCT to CAP is a big impact to teaching pedagogy as well as material development and selection.

BCT and CAP include two sections: single multiple choice questions and group questions, and the test type in both exams is multiple-choice question which is accompanied with four options. Group questions include two parts: reading comprehension test and cloze test. The section of group questions puts focus on evaluating the test takers' reading ability. The reading articles in group questions are authentic and multi-constructed with dialogues, descriptions, poems and comics with different genres.

From the sample questions in the past exams in CAP, we find that the reading passages were longer than BCT and the target cloze items were not a word or phrase anymore. The options included more words or even a sentence. Therefore, the test takers need to have the ability of dealing with complicated information and inferring between sentences to decide which answer is correct. However, the advanced ability to comprehend the complicated information is not stated in the competence indicators of Grade 1-9 Curriculum Guidelines (Appendix B) at the stage of junior high school. On the contrary, this advanced ability meets the competence of new Grade 1-12 Curriculum Guidelines (Appendix C) which is going to implement in 2018.

The cloze test is used to evaluate learners' command of discourse structure or linguistic knowledge (Bachman 1982). The evaluation of discourse structure puts focus on the ability of searching for the correct answer beyond the sentential level to make the text coherent. The linguistic knowledge emphasizes the grammatical points of tense forms, referential determiners, and other syntactic elements. The cloze test thus is an integrative test because it involves several types of linguistic knowledge in making the text logical and reasonable. For different age of language learners, cloze test is one of the most difficult parts in reading test because it targets at integrated comprehension skills. Besides, it is used widely in many testing contexts in Taiwan

which is from high-stakes exams to quizzes in class. Learners must pay more attention on the cloze items; otherwise, they cannot handle the complex tasks successfully.

In Taiwan, some research analyzes cloze test in senior high schools. For the characteristics of the cloze procedure, Yang (1996) investigates what the principles the local teachers used to generate the cloze items in monthly exams. For the analysis of cloze items, Cheng (2007) categorizes cloze items of the main entrance exams to college, Scholastic Aptitude English Test (SAET, hereafter) and Department Required English Test (DRET, hereafter), into five cohesion types according to the classification of Halliday and Hasan (1967). Chou (2009) utilizes Jonz's (1990) taxonomy and a modified classification procedure to categorize items in a section called "Zong-He-Ce-Yian" as "cloze test" in SAET and DRET. However, little research has been conducted in the field on the cloze test at the stage of junior high school even though the cloze test is a crucial testing type whether in formal entrance exam or monthly exams at school.

The new test type of senior high school entrance exam, CAP, differs greatly from the previous test. The new test type is closely examined by teachers, parents and students because it is a high-stakes test. The change shows an important perspective for the teachers in junior high school to analyze the language knowledge involved in cloze tests. The present study is thus motivated to investigate the cloze test for the better understanding of CAP, which in turn may cast light on teaching and learning.

The major research questions addressed in the present study are as follows:

1. What kind of language knowledge is measured by cloze items of BCT and CAP from 2010 to 2017?
2. What are the differences between cloze items of BCT and CAP from 2010 to 2017?

### **Significance of the Study**

Since the cloze test needs more integrated language ability for students, the significance of the present study can be illustrated in three aspects. The first is to discover the language knowledge on cloze test in senior high school entrance exams from 2010 to 2017. Secondly, the study aims at the characteristics of cloze items in both tests. The differences of cloze items between BCT and CAP will be compared in the present study. Lastly, for pedagogical application, since most teachers are more familiar with BCT and is still figuring out the content of CAP, the present study can give teachers some suggestions for their teaching activities in junior high school, and further design different lessons to address the important language knowledge to enhance assistance in students' language learning.



## **CHAPTER TWO**

### **LITERATURE REVIEW**

In this section, literature review on this topic is further divided into five parts. The first part is the relationship between testing and assessment in the curriculum design and approaches to language testing. The second part is the development of cloze test and the elaboration of four major types of cloze test. The third part is about what kind of language skill cloze tests measures. The fourth part is about the research on cloze test in Taiwan. The last part is the introduction of BCT and CAP.

#### **Testing and Assessment in Curriculum**

The well-constructed curriculum offers a proper framework to promote learning efficiency. According to Nicholls and Nicholls (1972), the stages of curriculum design are (1) setting objectives, (2) devising methods and choosing materials, (3) conducting assessment and (4) giving feedback to the objectives. In order to achieve the defined goals, each stage interacts with one another. The curriculum design should be a continuous cyclical process, which needs to be planned and modified constantly based on feedbacks from assessments (Tyler, 2010).

The role of testing and assessment is one major element for constructing an appropriate curriculum in teaching. In order to inspect students' learning as well as teachers' curriculum design, testing thus provides feedbacks to teachers for evaluating the effectiveness in teaching, examining the appropriateness of the textbook and reflecting the effectiveness of the syllabus. The test result thus provides clear information for teachers or designers to re-evaluate the objectives, teaching materials or teaching methods to modify the curricula.

The relationship between curriculum and assessment is dynamic and interactive, and the assessment should not be parted from the language learning process (Brown, 1995; Brown, 2005: 252; Dick, Carey & Carey, 2002: 2). Hughes (2003) stated that an

accurate measurement of test takers' target skill is essential to a language classroom, as it provides formative and summative information about the learning progress and achievement, which in turn gives the instructor beneficial feedback in how and what to teach.

The selection of test depends on different curriculum purpose. For inspecting students' progress in learning, formative assessment is a proper choice to evaluate their weakness and strength. Aside from providing the feedback for curriculum modification, student can also monitor their learning progress by formative assessment. Summative assessment, on the other hand, is used to measure what has been achieved by the end of a semester or a year which is called formal test (Hughes,2003).

### *Approaches to Language Testing*

From the chronological order, language testing can be classified into four main approaches:(1) the essay-translation approach, (2) the structuralist approach, (3) the integrative approach and (4) the communicative approach (Heaton, 1975).

The essay-translation approach is commonly used at pre-scientific stage in language testing. The tests consist of essay writing, translation and grammatical analysis. The approach can apply to examinees of any level. However, the tests have a heavy literary and culture bias as well as the teachers' subjective judgment.

On the contrast, the behaviorism teaching approach during the period of the 1940s and 1950s is associated with the psychometric-structuralist in language testing. The structuralist approach views language learning as the systematic acquisition of a set of habits. It draws attention on a set of structure linguistics: phonology, vocabulary and grammar. The tests of four abilities (reading, writing, listening and speaking) are also separated from one another because testing one thing at a time is emphasized in



this approach. The approach in language testing put focus on the mastery in separate elements or skills instead of integrated ability.

However, having a good command of using language in a larger context cannot be measured through discrete-point test. The students are not able to have ability to communicate outside the classroom in the real world (Richards & Rodgers, 2003). Furthermore, it is found that there is no positive relationship between learners' performances on discrete-point test and their language ability in communication (Richards & Rodgers, 2003). Hence, in the 1970s and the 1980s, new language teaching methods start growing, and put emphasis on integrative language ability.

This new era of language teaching method is called integrative-sociolinguistic period (Spolsky, 1996). Starting from the integrative approach, providing a real context for the communication purpose is the main characteristics in this period. The focus of testing shifted on the testing of language in context with communicative effect. The focus is different from the previous approach that put emphasis on a discrete-point item. A discrete-point test focuses on one grammatical point at a time, whereas the integrative test measures more than one point at a time. (Oller, 1979; Hadley, 2000; Hughes, 2003).

In this approach, the language testing techniques are designed to increase learners' contextualized knowledge. The test designers aim to test learner's ability by utilizing different aspects of language knowledge instead of isolating pieces. The testing formats also followed this principle to change. Many scholars pointed out the cloze tests and dictations were considered as the two types of integrative tests (Oller, 1979; Bailey, 1998; Brown, 2001; Brown, 2005).

Lastly, the communicative approach is focus on how language is performed in communication. In this approach, most tasks are related to students' real life, and the effectiveness of communication is the target purpose instead of the language forms.

The measurement of communicative approach is based on the concept of divisibility hypothesis in language. The learners thus have several profiles to measure the different language performances.

### **The Development of Cloze Test**

This section addressed the development of the cloze test, including three parts: measurement of the readability, measurement of learners' reading comprehension, and types of modified cloze test.

#### *Measurement of the Readability*

The readability of a text affects the difficulty of a test. The readability means the difficulty level of a reading passage. The readability of a passage can be measured by sentence length, the complexity of sentences, vocabulary difficulty, and others. Dale and Chall (1948) and the Flesh (1948) are the two most widely used formulas. However, Taylor (1953) proposed the cloze procedure because he thought that the previous two formulas used to measure the readability of a prose that dealt with some superficial aspects of the text. The cloze procedure integrates most of the potential elements that may influence the readability of a prose.

The invention of the cloze procedure opens a crucial page of language testing, that is, the total language context. The total language context indicates that "the ability to identify, learn and produce any language symbols depends heavily on the variable degrees to which it is associated with everything else by larger and meaningful overall combinations " (Taylor, 1953, p. 418).

The performance in a cloze passage represents learners' understanding of the text. Oller (1979) suggests that teachers can select the appropriate level of cloze passages for learners based on their current level. Generally, two common methods are used to score a cloze passage which represent learners' understanding of a specific passage: The exact word method and the acceptable word method. The exact word method is

that the test takers need to provide the exact word as the original word from the original passage. Conversely, the acceptable word method indicates that any response which is grammatically correct and make sense in the context is regarded as correct answers.

### *Measurement of the Reading Comprehension*

Reading comprehension is treated as an important ability in language learning. The process of reading comprehension covers all the meaningful elements of a texts that are integrated into a coherent whole during text process (Kintch and van Dijk ,1978). The reading ability involves organizing and arranging the written information to get the whole pictures of the texts. The reading comprehension occurs when the readers make sense out of passages, and integrate separately presented ideas into a unitary whole (Koler, 1973; Wittrock, 1975). Reading comprehension test is used widely from class quizzes to high-stakes examination. There are numerous ways of testing reading comprehension, ranging from multiple-choice items to open-ended questions (Heaton, 1998).

Since the result of the cloze test can be utilized as the measurement to present learners' understanding of a passage, researchers have been interested in the relationship between cloze test and reading comprehension test.

Bormuth (1969) conducted a study for 4<sup>th</sup> to 6<sup>th</sup> grade students. The nine reading passages with readability index between 4.5 to 6.6. of Dale & Chall's formula were adopted and modified into a cloze test. At the same time, seven multiple choices are provided to measure learners' reading comprehension. The study shows the results of the cloze test are highly related with the reading comprehension test.

Dupuis (1980) also have a similar result for the tenth-grade students (senior high school level) in the study. Dupuis intended to investigate whether cloze procedure is also an effective instrument in measuring learners' comprehension in different genres.

She selected two short stories and one novel, and adopted the cloze test by every  $n^{\text{th}}$  word deletion to construct a cloze passage. Each passage was modified into a cloze passage accompanied with some reading comprehension questions. The study reveals a highly-connected relationship between the cloze test and the reading comprehension test in both genres.

In Taiwan, Huang (1992) conducted a study to compare the result of the two types cloze tests with reading comprehension in TOEFL. The cloze procedure of two types cloze tests in this study were the C-test and multiple choice. The study revealed that both cloze tests effectively predicted learners' performance in reading comprehension.

The relationship between cloze test and reading comprehension test is highly related as shown in the above studies. The similar conclusion provides a notion that cloze tests can be used to test both micro and macro-level language skills. However, in Greene's study (2001) of evaluating the reasoning ability by a cloze passage with deleted discourse markers, he concluded that cloze test is a reliable tool to evaluate learners' reasoning ability of a passage when the focus is on the connection between sentences, but it cannot be used to evaluate learners content knowledge in reading comprehension of a passage. Alderson (2005) made a similar conclusion that cloze procedure is word-based, it thus may not evaluate some reading skills.

#### *Types of Modified Cloze Tests*

Bailey (1998) pointed out "the process of deleting word to create a cloze is known as mutilation" (p.62). According to the way how a cloze item is generated, four general types of cloze tests are commonly used: fixed-ratio deletion, the C-test, the rational deletion and multiple choice. The following section will present the introduction of the four general types of the cloze test.

The first type of cloze test is called fixed-ratio deletion ( $n^{th}$  word deletion). Every  $n^{th}$  word is deleted no matter what word it is in a selected passage. The procedure tends to test various items types, including the governed constraints in local grammatical knowledge and long-range textual ones. The advantage of the fixed ratio deletion is that the test constructors can easily and objectively generate cloze items; however, the items may be either too difficult or meaningless to the learners. Also, the different deletion frequencies ( $6^{th}$ ,  $8^{th}$  or  $10^{th}$ ) of the same passage may produce different difficulty levels of the cloze test (Alderson, 1979, 1980).

The second type is C-test. Two types of C-test have been identified. The first type of C-test is a type of language test in which the students read a brief paragraph in the target language with the first two sentences are left intact. Starting from the third sentence, the individual letters of every other word are replaced by the blanks (Bailey, 1998). Another type of C-test keeps the first sentence remain intact only, and the second half of every second word is deleted from the second sentence. The advantages of the C-test are that only exact scoring is necessary and it possibly covers a wider range of topics, styles, and levels of ability. On the other hand, the disadvantage of C-test is that it is harder to read than a cloze passage, and the correct answers can often be found in the surrounding text (Hughes, 2003).

The third type is the rational deletion that is also called selective deletion, which means the words are deleted based on some rational decisions. Since the fixed-ratio deletion does not test on target item in language, the rational deletion suggests a new test type for test developers to control over the types of words deletion to measure the target language ability. Rational cloze tests are as "professionally developed" tests in general intend to measure high-order reading processing by fine-tuning the items (Bensoussan, 1990; Storey, 1997).

Bachman's study (1982) suggested that "cloze tests are not necessarily monotonic tests of general and specific traits." The result showed that rational cloze was capable of measuring higher order skills: cohesion and coherence in the text.

In Bachman's study in 1985, he conducted another study about the proportions of different cloze items types deleted by fixed-ratio deletion and rational deletion.

Bachman developed a classification framework for cloze item types according to the different required level of the context to fill in the blank, which are (1) Within Clause, (2) Across Clause, Within Sentence, (3) Across Sentence, Within Text, and (4) Extratextual. The result revealed that fixed-ratio deletion tends to produce items of (1) *Within Clause* and (4) *Extratextual*. The two types refer to the blanks that can be filled in simply by using clause-level grammatical knowledge or world /extra textual knowledge. Contrast to fixed-ratio deletion, the rational deletion produces more items of (2) *Across Clause, Within Sentence*, and (3) *Across Sentence, Within Text*. Types (2) and (3) indicate that the test-takers are required to fill in the items by using the clues provided by the text beyond clause or sentence.

A similar result is found in Jonz's study in 1990. He applied Bachman's (1985) four-category analytical approach on seven cloze passages. Figure 2.1 presents the flowchart of Jonz's interpretation about Bachman's four categories.

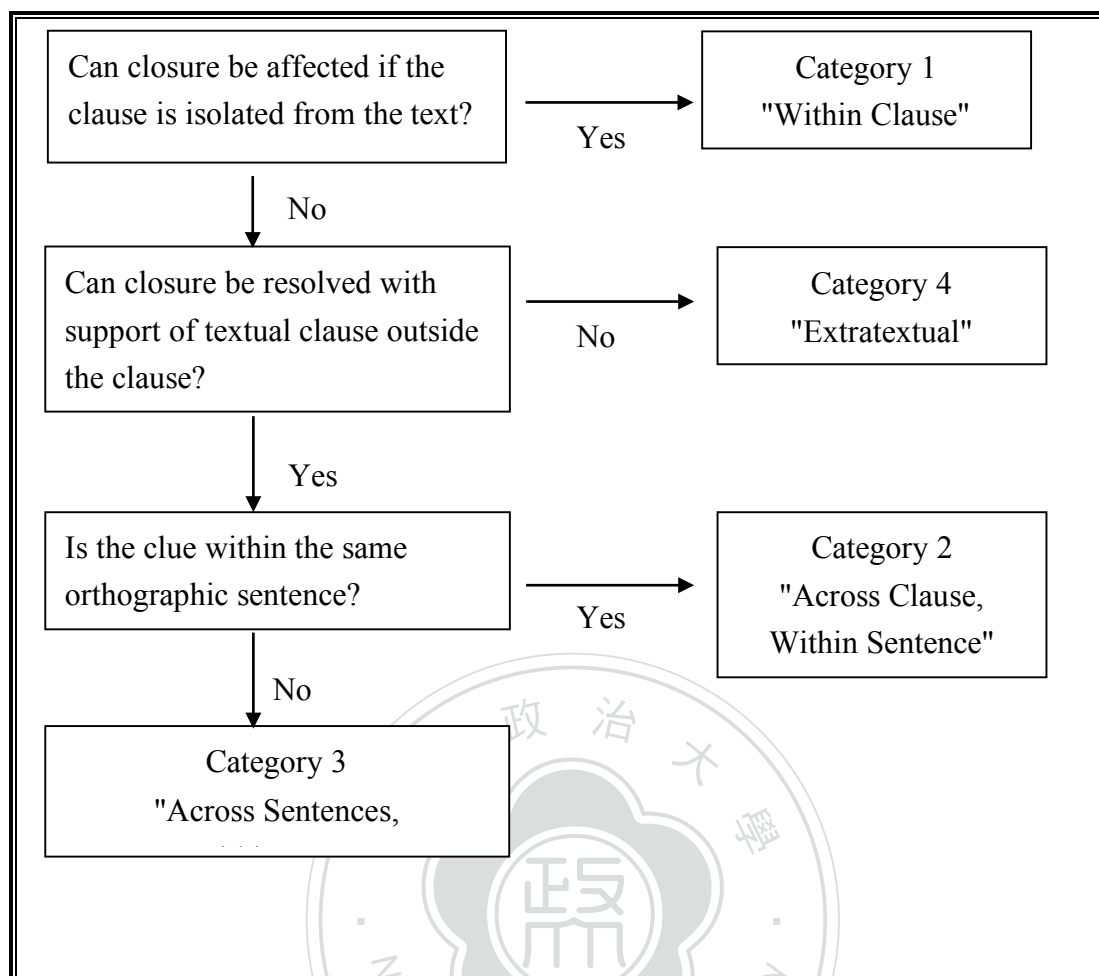


Figure 2.1 The Flowchart of Jonz's Interpretation of Bachman's Categories

He further divided the Category 1 - *Within Clause* into two subcategories: (a) Clause-level lexis, and (b) Clause-level syntax. He proposed that the division is to uncover the evidence that cloze tests measure both the high-order (lexical choice) and low-order (syntactic knowledge) comprehension. Besides, he proposed a classification key of the subtypes under the five categories. The classification key provides a precise source to classify the cloze items. The present study adopts Jonz's framework to analysis the cloze items in BCT and CAP from 2010 to 2017, and the complete classification will be presented in chapter 3.

The previous three types of cloze test require test-takers to fill in the blanks, in terms of scoring, two methods are adopted: the exact word and the acceptable-word. Comparing these two methods, the exact word is easier for scoring; however, it may

not actually evaluate student's performance in using the target language. On the other hand, the acceptable-word method provides more opportunities for learners' creativity in using the language, but it increases the difficulty and time of scoring. From the disadvantages of above scoring methods, multiple-choice cloze test becomes a popular test format, that is each cloze item is followed by four alternative options, in many formal and large-scale examinations because of the easy scoring.

The fourth modified test type is multiple-choice cloze test. Panackal and Heft (1978) compared learners' performance on Scholastic Aptitude Test (SAT, hereafter) by open-ended cloze test and multiple-choice cloze test. The result reveals that both test types are comparable to SAT, and multiple-choice cloze test can be used to replace the open-ended cloze test.

Hinofotis and Snow (1978) used 25 multiple choice items and 25 open-ended cloze test items to investigate whether multiple choice cloze test work equally well as open-ended cloze test. The study showed that multiple choice items are easier in terms of item facility indices, and open-ended cloze test is better in terms of item discrimination indices. They concluded that even there are different performances on both tests, the multiple choice is a more promising evaluative tool because of ease of scoring.

Although the multiple choice cloze test is relatively easy and convenient in scoring, constructing a multiple choice cloze test is more complicated than constructing an open-ended cloze test (Hinofotis & Snow, 1978; Wu, 2002). Wu (2002) pointed out the deletion and alternatives have an effect on the test, and the limitation of the alternatives affect test takers' comprehension processing. In order to construct a multiple choice cloze test with validity, it is necessary to check item facility and discrimination indices and carry out the pretest (Hinofotis & Snow, 1978). By doing so, the improper cloze items should be discriminated and modified.



### **What Kind of Language Skill Does Cloze Measure?**

Cloze tests are widely used for measuring text comprehension (Sharp, 2009; Schmitt & Sha, 2009; Miller, Dewitt, McCleary & O'Keefe, 2009), proficiency (Keshavarz & Salimi, 2007), collocational knowledge (Stuart & Eve, 2009; Keshavarz & Salimi, 2007), as a test of translation (Ito, 2004), and a pedagogical tool (Dastjerdi & Talebinezhad, 2006; Lee, 2008).

The cloze procedure has been long studied in the divergent opinion about what kind of language ability the cloze procedure measure; whether it measures learners' higher or lower-order skills. Some argued learner's overall language proficiency is able to be tested and obtained from the learning result by the cloze test (Chavez-Oller et al., 1994; Bachman, 1982; Sasaki, 2000). On the other hand, some argued the cloze test can only test local language skills because of the insensitivity (Alderson, 1979; Shanahan, 1982; Markman, 1985).

Chavez-Oller et al. (1994) found that the cloze items are sensitive to long-range constraints, and the constraints across sentence boundaries facilitate performances on some cloze items. The items which refer to the clues beyond sentential level can discriminate high achievers from low achievers.

Bachman (1982) proposes that the cloze test items can be classified into three traits: syntactic, cohesive and strategic. The syntactic trait which depended on clause-level. The cohesive trait means the closure is on interclausal and intersentential cohesive structure (as described by Halliday and Hasan in 1976). The strategic trait refers to "long-range" patterns of coherence. The study indicates that the rational deletion can be used to test on specific abilities and measure the higher-order skills.

Sasaki (2000) investigated the effects of cultural schemata on test-taking processes of cloze tests and used immediate retrospective introspection as one of the

data sources. The study revealed that participants did use information beyond the sentence level to answer.

However, some studies suggested that cloze test is not able to test learners' integrative ability, it only tests local syntactic skills. The changes of the deletion rates also change the measurement in different abilities. Shanahan et al. (1982) conducted a study to examine whether the cloze passages are able to measure learners' ability from the information across sentence boundaries. The researchers deleted every 5th word in the passage, and used sequential and scramble cloze passages. Besides, they added one more type of cloze passage in the study: they inserted a single sentence into the original cloze passage. They assumed that the learners could not use the information across sentence boundaries under the three types of cloze passages. The result revealed that there are no differences in learners' performances. The test takers who have only a single sentence also have good performances like those have the same sentence embedded in meaningful context.

Markman (1985) adopted the fixed ratio deletion to examine the intersentential sensitivity of the rational deletion cloze procedure. He deleted every 7<sup>th</sup> and 9<sup>th</sup> word in both sequential and scramble texts, and placed adjacent sentences in nonadjacent position to destroy the intersentential ties in scramble texts. The result shows there is no differences in performance on both cloze tasks.

A similar conclusion in the above studies claim that the cloze procedure measure the lower-order skills. In the fixed-ratio deletion, the test constructors are not able to control the test items in specific abilities. In this way, only the cloze items at clausal and sentential level can be generate. Therefore, Alderson (1980) suggested that if test constructors carefully design the deletion closure and make sure the clues for the cloze items are beyond sentences, the cloze test is capable of measuring language knowledge across sentence boundaries.

### **Relevant Research on Cloze Test in Taiwan**

The cloze test has been adopted in various contexts, which is from monthly exams to high-stake tests, and some research addressed the topic on cloze tests in Taiwan (Yang,1996; Huang, 1997; Change, 1994; Chou,2009 ).

Yang (1996) analyzed the cloze items in monthly exams of secondary school, and interviewed with local teachers about the principles they use to generate cloze items. He indicated that the cloze tests have often been misused in high schools in Taiwan because "most teachers are not aware of the true characteristics of the cloze procedure and tend to interpret it as gap-filling in a continued discourse without concerning themselves with contextual dependence of test items" (p.62). In Yang's study, he found that many multiple-choice cloze tests tend to serve as discrete point test rather than integrative test because too many local items are tested instead of cross-sentential contextual clues.

Huang (1997) studied the cloze test items of Joint College Entrance Exam (JCEE) from the year 1982 to 1993 by a qualitative analysis. He divided cloze items into two types: global and local items. The local item means the item could be restored within clause or sentence level. The global item indicates that the contextual clues for the cloze item are beyond sentence. The result revealed that global items only accounted for 40 (16.33%) out 245 items. JCEE put much emphasis on local items, and the modified cloze test is regarded as discrete-point test. Change (1994) found the similar conclusion that the cloze items in Taiwan focus more on local items. He suggested test constructors should balance both local and global items to measure learners' overall language ability.

Chou (2009) applied Jonz's taxonomy to evaluate the cloze items of the two college entrance exams in the section of Zong-He-Ce-Yian in SAET and DRET from 2004 to 2008. The finding of the study reveals that 78% and 69% of the

Zong-He-Ce-Yian in SAET and DRET are classified as local items, and 22% and 31% in SAET and DRET are classified as global items (pp.79-80). These two nation-wide exams tend to measure examinees' discrete-point skills rather than integrative skills (p.85).

To balance the cloze items on local and global skills in cloze test, test constructors can use some criteria to check. Wu (2002) classified item generation approaches into three types: syntactic-, lexical, and discourse-oriented approaches. The previous two approaches are considered as items for local skills, and discourse-oriented approach is considered as global items.

Besides, Bachman's taxonomy (1985) can be used for evaluating whether a cloze item is discrete-point oriented or integrative oriented. Yang (1996) pointed out that (1) *Within clause* and (4) *Extratextual* are associated with discrete-point skills because (1) *Within clause* deals with contextual clues only within clause level, and (4) *Extratextual* requires the knowledge without inferring from the text. Types (2) *Across Clause*, *Within Sentence* and (3) *Across Sentence*, *Within Text* deal with interclausal or intersentential integrative skills. In most local context, the multiple-choice cloze test and the procedure of rational deletion are adopted, the test constructors should be more aware of what kind of the language knowledge in cloze items aim to measure as well as the balanced distribution of the measurement in global and local language knowledge.

### **BCT and CAP**

In Taiwan, the main standardized entrance exam of senior high school for ninth-grade students from 1999 to 2012 was BCT. Starting from 2014, CAP is the new test type. The goals of academic attainment testing construction of BCT and CAP are according to the competence indicators in Grade 1-9 Curriculum Guidelines which is related to the English learning objectives within 7<sup>th</sup> to 9<sup>th</sup> grade. The format of

reading test in these two exams consists of two parts: single multiple choice questions and group questions. Group questions include reading comprehension test and cloze test. The cloze test is conducted by rational deletion in multiple choice questions with four options in both exams. The test genres for cloze test include written passages, dialogues, ads, chats, and other authentic ones.

### *BCT*

BCT was a main entrance exam of senior high school from 1999 to 2013 for junior graduates. From 1999 to 2011, BCT had been conducted twice a year in May and July. Starting from 2012 to 2013, BCT was changed to be conducted only once in a year in June. BCT is designed at the level of easy according to the official announcement of Research Center for Psychological and Educational Testing. The result of the test performance is counted in the way of norm-reference and presented by the scale scores. The total scale scores of BCT is 412 which include 80 scores in each five subjects from 2009 (before 2009, 60 scores in each subjects) and 12 scores in composition from 2007.

There are 45 test questions in total for 70 minutes. The design of the item types in BCT are the multiple-choice questions with four options. There are 10 or 11 group questions designed in the standardized exam. The group questions are arranged by reading comprehension test and cloze test randomly. The number of the cloze test of group questions are mainly constructed by 2 passages with totally 6 or 7 cloze items in the test each year.

### *CAP*

In 2014, BCT was replaced by the new test type CAP. Before the formal practice of CAP, a sample test was conducted in 2013. According to RCPET, the objectives of the CAP are as follows: (1) CAP mainly focuses on basic, key and important concepts

of English, (2) the various genres and topics are included; (3) the emphasis is focus on natural English language usage instead of complicated and fragmental memory.

For the degree of difficulty, CAP is at the level of appropriate which means the proportion of four difficulty levels is equally distributed among the test. The result of the test performance is counted in the way of criterion-reference and presented by three levels: mastery (A++, A+ and A), basic (B++, B+ and B) and improvement needed (C).

Besides, the English test contains two parts in CAP, listening test and reading test, with a total 60 of multiple choice questions. From the sample test in 2013 to formal test in 2014, these two parts are combined into one same exam paper, and the test time is 80 minutes. However, starting from 2015, the listening test and reading test are separated into two independent test sections which are presented in two different exam papers. The test time for listening test is 25 minutes for 20 multiple choice questions with three options, and for the reading test is 60 minutes for 40 multiple choice questions with four options.

Above all, there are some changes from BCT and CAP, and the differences certainly bring some impacts on teaching and learning. The current study is thus motivated to investigate what kind of language knowledge of both exams aim to measure by adopting Jonz's taxonomy. Furthermore, the current study also analyzes the differences of items types in the cloze tests of BCT and CAP from 2010 to 2017.

## **CHAPTER THREE**

### **METHODOLOGY**

The present study aims to analyze and observe what kind of the language abilities measured in the cloze test section of the two national senior high school entrance exams in Taiwan, BCT and CAP. Jonz' taxonomy (1990) is adopted as the instrument to classified the two types of materials into five categories, and then each cloze item is classified into one subtype under each category according to their test targets. In this chapter, the materials, instrument and procedure of data analysis are presented.

#### **Materials**

The materials in the present study are collected from BCT and CAP. The test items are obtained from the website of Comprehensive Assessment Program for Junior High School Students (<http://cap.ntnu.edu.tw>). The data of both exams include the test items in the past 8 years. The current study will analyze the cloze items from 2010 to 2013 in BCT, and analyze the cloze items from 2014 to 2017 in CAP. Table 3.1 presents the number of reading passages and the number of test items in BCT and CAP from 2010 to 2017.

Table 3. 1 Distribution of Reading Passages and Test Items in Cloze Test

Data Year (1st or 2nd)	Number of Reading Passages	Number of Test Items
2010-1 BCT	2	6
2010-2 BCT	2	7
2011-1 BCT	2	6
2011-2 BCT	2	7
2012 BCT	2	6
2013 BCT	2	6
2014 CAP	2	7
2015 CAP	2	6
2016 CAP	2	7
2017 CAP	2	8
Total	20	66

### Instrument

The present study aims to find what kind of language knowledge is measured in the standardized national entrance exams of senior high school in Taiwan. The researcher applies the taxonomy in Jonz's study in 1990 to analyze the features of cloze items in BCT and CAP.

In 1990, Jonz analyzes cloze passages based on Bachman's framework (1985) to categorize cloze items into four types according to the level of context required for recovering the missing word. However, Bachman's study does not provide further explanation for how and why each item is assigned to a specific category. Therefore, Jonz develops his operationalization rationale for the interpretation of cloze deletions of Bachman's categories. To have a deeper understanding of the test items, Jonz (1990) further divides two levels under Category 1-Within Clause: Clause-level syntax and Clause-level lexis. Jonz proposes an assumption that the lexical choice represents the high-order knowledge, and the syntactic knowledge represents the low-order one (p.70).



The present study adopts Jonz's model that includes two levels in *Category 1* to collect more data for better understanding of the cloze items in local context. Hence, each cloze test item will be assigned into 5 categories in the current study:

(1) Clause-Syntax, (2) Clause-Lexis, (3) Across Clause, Within Sentence, (4) Across Sentence, Within Text, and (5) Extratextual (C1, C2, C3, C4, C5 hereafter).

Furthermore, Jonz lists several subtypes of grammatical structure under the five categories. The subtypes under each category provide not only the more precise source to classify the language elements that the cloze items target at but the better understanding on what kind of language knowledge is measured (pp.76-77). Table 3.2 presents the categorization of each subtype provided by Jonz.

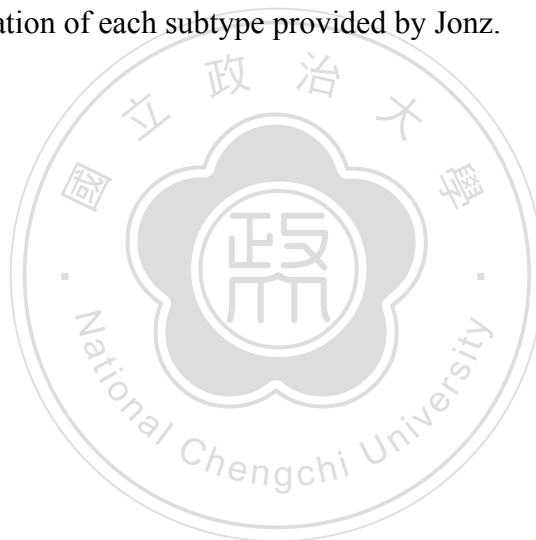


Table 3. 2 Jonz's (1990) Subtypes of Grammatical Points Tested in Cloze Items (pp.76-77)

Clause-level Syntax	
1-S-a	determiner without text antecedent
1-S-b	conjunction below clause level
1-S-c	role-marking preposition
1-S-d	relative word in headless clause
1-S-e	Complementizer
1-S-f	auxiliary
1-S-g	negation
1-S-h	copula BE & other intensive relationships
1-S-i	syntactic <i>there</i> or syntactic <i>it</i> + Be
1-S-j	syntactic elements in cleft or pseudocleft
Clause-level Lexis	
1-L-a	referential determiner or deictic within clause
1-L-b	pre- & post determiners
1-L-c	pronoun (coreferent with clause)
1-L-d	indefinite proform (including <i>you</i> as indefinite)
1-L-e	preposition other than role-marker
1-L-f	collocation with clause
1-L-g	reiteration with clause
1-L-h	multipart item
1-L-i	qualifier in AdjP & AdvP
Across Clause, Within Sentence	
2-a	referential determiner or deictic across clauses
2-b	conjunction at clause level
2-c	pronoun (coreferent across clauses)
2-d	relative word in headed clause
2-e	collocation across clauses
2-f	reiteration across clauses
2-g	subordinator
2-h	<i>it</i> of extraposition
Across Sentence, Within Text	
3-a	referential determiner or deictic across sentences
3-b	conjunction between sentences
3-c	pronoun (coreferent across sentences)
3-d	collocation across sentences

3-e	reiteration across sentences
3-f	textual discourse marker
Extratextual	
4-a	speech-role pronoun without textual antecedent
4-b	lexical item with little or no local clue

Most subtypes in Jonz's taxonomy are covered in grammar books, like pronoun, auxiliary and negation. The following sections are the further explanations of some subtypes that students are not so familiar with when they are learning English, which are: *1-S-e complementizer*, *1-S-h copula BE & other intensive relationships*, *1-L-h multipart item*, *2-h it of extraposition*, and *3-f textual discourse marker*.

#### *1-S-e: complementizer*

The complementizer within a clause includes two types: subject complement and object complement. Subject complement is a word or phrase that follows a linking verb to describe or identify the subject. The linking verbs are like “*am, is, are, appear, look, smell, and others.*” A subject complement could be an adjective, a noun, an adverb, a pronoun, or others.

(1) That pie looks delicious.

(2) Mosquitoes are everywhere.

In Example 1, “*delicious*” is the complementizer to describe how that pie looks. In Example 2, “*everywhere*” is an adverb that identifies where the mosquitoes are. On the other hand, object complement is utilized to describe or identify the object.

(3) The President appointed her Secretary of State.

In Example 3, “*her*” is the object of the transitive verb “*appoint.*” In addition to “*her*”, we need a word to complete the meaning of the verb. The noun “*Secretary of State*” completes the meaning of the clause and functions as a complement to the object “*her.*”

### *1-S-h: copula BE & other intensive relationships*

The intensive relationship means to focus on the subject, and the following intensive verbs serve the function as the subject complement.

(4) Jenny is a doctor.

(5) Jason looks very young for his age.

In the above examples 4 and 5, the usage of *copula be* and *intensive relationship*, such as “*am*” and “*look*” function as the join of an adjective or a noun to each subject.

In Example 4, the focus of the sentence is Jenny and what she is. In Example 5, the focus of the sentence is Jason and how he looks. Copular verbs, such as “*taste, look, smell*” also known as the linking verb to play the role of subject complement.

### *1-L-h: multipart item*

A multipart item refers to the items which contain more than one, such as phrasal verbs, “*make up, turn on, take off.*” A phrasal verb includes a simple verb plus a particle, or in rare cases two particles, but it must be considered a single lexical item.

(6) He took off his rain coat.

(7) The plane is going to take off in 5 minutes.

(8) He took two weeks off in December to have a vacation in France.

In the above three examples, each sentence contains the same phrase verb “*take off*”; however, there are three different meanings that learners should comprehend depend on the other components in the sentence.

### *2-h: it of extraposition*

This subtype refers to a construction in which a clause that acts as a subject is moved or extraposed to the end of the sentence, and replaced by dummy “*it*” in the original position.

(9) It is true that you have been misled.

(10) It is exciting to watch the baseball game at baseball field in person.

The subjects in examples 9 and 10 are replaced by dummy “*It*”, and the following “*that- clause*” and infinitive “*to*” bring out the complete meaning of the two subjects. By doing so, the writer can avoid a problem of having a heavy subject in writing.

### *3-f: textual discourse marker*

Discourse markers are words or phrases used to show how discourse is constructed. They connect and organize what we say or write, or to express attitude.

(11) A: I like Maggie. She is adorable. B: So, you’ll ask her out then.

(12) My father promised to have a family trip this summer. However, he changed his mind last week.

In Example 11, the expression “*So*” shows what is said is the result of what was said before. It serves the function as a logical consequence. In Example 12, the discourse marker “*however*” is used to begin a new sentence which contracts two ideas.

The other subtypes of Jonz’ taxonomy are covered by grammar courses at the stage of junior high school; for example, determiners “*the*” and “*that*”, conjunction “*or*” and “*but*”, negation “*no*” and “*not*”. In addition, learners are also required to master the principles of basic tenses from English courses at this stage; for instance, present simple tense, past tense, future tense, present and past progressive, and present and past perfect tense.

### **Modification of Jonz's Subcategories**

Jonz (1990) analyzes eight cloze passages that are developed either by fixed-ratio or rational deletion. When the cloze passages are applied on the fixed-ratio deletion, every  $n^{\text{th}}$  word is deleted. However, BCT and CAP adopt the procedure of rational deletion which is different from Jonz’s. As a result, it may cast a problem that not each cloze item of BCT and CAP can be classified properly into Jonz’s taxonomy.

The rational deletion is adopted in BCT and CAP, and the cloze tests in both exams are designed as multiple choice questions, that is, the test-takers must choose

the correct answer from four options. Therefore, the test developers are able to control the types and elements of the word/ sentence deletion to test the target points.

On the contrary, the standard cloze procedure requires the test-takers to fill in the blanks by themselves. As the case stands, some cloze items in Jonz's subtypes cannot be found in BCT and CAP if they do not represent critical language knowledge. The test constructors in local context may also delete more than one word or even a complete sentence in an item. Therefore, the researcher makes six modifications from Jonz's taxonomy in the present study.

The first modification is to distinguish 2-*b: conjunction* at clause level from 2-*g: subordinator*. In general, conjunction includes three subtypes: coordinating, subordinating and correlative. Coordinating conjunctions connect two equal units of a sentence together, such as “*and, but, nor, yet*”. On the other hand, subordinating conjunctions connect two unequal parts of a sentence: main clause and subordinating clause. The subordinating conjunctions are like “*before, after, when, because*.” The following example shows two parts of a sentence.

(13) I got wet because I didn't bring an umbrella with me this morning.

In Example 13, “*I got wet*” is the main clause and “*because*” connects the subordinating clause “*I didn't bring an umbrella with me this morning*.” Two unequal parts conveys the relation of cause and effect.

The third one is correlative conjunctions which are always used in pairs, such like “*both/and, neither/ nor* and “*not only/ but also*”, often tested at graders 7 to 9. They join similar important sentence elements, so they are classified into coordinating conjunction in the current study. Therefore, 2-*b: conjunction at clause level* is modified as 2-*b: coordinator*, and 2-*g: subordinator* remains the same.

The second modification is in *Category 5: Extratextual*. According to Jonz's taxonomy, *Category 5* is known as *Extratextual*, and two subtypes are provided under

this category: (1) *speech-role pronoun without textual antecedent*, (2) *lexical item with little or no local clue*.

(14) He had to do a lot of \_THINGS\_ he didn't like to do \_JUST\_ because he was going to go away to college.

(From Jonz, 1990, p.78)

In Example 14, the test takers need to fill up the blanks, and the test item is categorized under the second type, indicating lexical item with little or no clue. However, as the cloze procedure is applied as multiple choices in BCT and CAP, the test-takers do not have to fill in the answers by themselves. Therefore, the adjustment in the category of *Extratextual* of Jonz's study in the standardized cloze procedure has to be made.

For the SLA learners, in order to pursue the goal of authenticity in language testing, some test formats like charts, ads, time tables, and maps in daily life are also provided for cloze tests design. The test-takers need to acquire the ability of understanding the maps to find out the correct direction, or have the ability of taking an order from a menu at a restaurant. One example of ads from BCT is presented as follows:

Mr. Lee: Wow! You look great in 25 glasses! Its glasses are very famous. How much is this pair?

Mr. Kao: Only 500 dollars.


Mr. Lee: Really? Where did you get the glasses?

Mr. Kao: At B. B. Department Store. It's having a sale these days.

Mr. Lee: Does the department store have any 26 on sale, too? I want to buy a pair for my wife's birthday.

Mr. Kao: Yes, it does. I remember I have the flyer. Here. You'll get a much better price if you buy three pairs at a time.

Mr. Lee: Thanks.

 flyer 廣告傳單

**B.B. DEPARTMENT STORE**  
Special Prices This Week

<p style="text-align: center;"><b>Glass House</b></p> <p style="text-align: center;">glass plate / dish / bowl any three pieces for \$1,500</p>	<p style="text-align: center;"><b>O.L.S. Glasses</b></p> <p style="text-align: center;">two pairs for \$500</p>
<p style="text-align: center;"><b>Rian's</b></p> <ul style="list-style-type: none"> <li>▪ glasses 1 for \$900</li> <li>▪ blue jeans 1 for \$800 3 for \$1,800</li> </ul>	<p style="text-align: center;"><b>NUMA</b></p> <ul style="list-style-type: none"> <li>* sports shoes               <ul style="list-style-type: none"> <li>- one pair for \$700</li> <li>- two pairs for \$1,200</li> </ul> </li> <li>* glasses \$500/pair</li> <li>* tennis shorts \$300/pair</li> </ul>

- (15) Mr. Lee: Wow! You look great in 25 glasses! Its glasses are very famous. How much is this pair?
- Mr. Kao: Only 500 dollars.
- Mr. Lee: Really? Where did you get the glasses?
- Mr. Kao: At B.B Department Store. It's having a sale these days.
- 25.(A) Glass House (B) O.L.S. (C) Rian's \*(D) NUMA  
(From BCT 2011-1st)

The cloze test in Example 15 is presented by a format of a written dialogue with extra textual of four ads. From the dialogue, there are two interlocutors discussing some shopping items and prizes. For Question 25, the clue provided is beyond the sentential level, that is, the examinees must understand the following sentences accompanied with the ads below to choose the correct answer. Therefore, when the genre design includes two parts: written passage and picture (ads, charts, maps and others), the category is categorized into the modified Category 5—*Extratextual*. In addition, there are only some key words in adds or charts, and they don't have the continuous text



and contextual clues. Hence, the subtype *4-b: lexical item with little or no local clue* is modified as *4-b: lexical item with little or no clues from the text* in the present study.

The third modification is to add one more subtype to *Category 1*. The application of tense is widely tested in local testing context which ranges from classroom quizzes to nationwide senior entrance examination in BCT and CAP. The usage of tense requires the test-takers to select the correct tense based on the clues provided within a clause or across clauses. Below is an example from BCT:

(16)

In the old times, people in the town \_18\_ their houses on both sides of the river. The houses in each block shared a public washing room. On Mondays, people came and washed their clothes together there. They hung their clothes at the back of each house.

18. (A) build (B) are building (C) have built \*(D) built

(From BCT 2010-2)

In Example 16, Question 18 is provided with the clue “*In the old times*” within the same clause. Based on the clue, test-takers are able to choose the most appropriate tense – past tense – for the item. In this case, the tense of the item is decided at the clause level of *Category 1: Clause – level Syntax*. However, there is no proper subtype among *Category 1* for the option of tense.

On the other hand, the lexis level part of *Category 2: I-L-a referential determiner or deictic within clause* involves the identification of the relation between elements within a clause, like the first speaker and the interlocutor, the time and the space (here and now). Therefore, *I-L-a referential determiner or deictic with clause*, is moved to *Category 1* under the level of syntax in the current study.

The fourth modification is to re-categorize about some subtypes *C1- syntax* and *C2 - lexis*. The subtypes which is associated to grammar knowledge like “*pronoun*” and “*referential determiner or deictic within clause*” are moved to C1. To have more

clear analysis of the cloze items in C2, the subtypes which relate to lexical usage, such as “*collocation*” and “*reiteration*” remain in the level of lexis.

The fifth modification is made in C4. In attempt to relate leaning to real communication, it is crucial for learners to have more appropriate language performances in authentic situation. The test constructors not only include written passages but daily dialogues in cloze tests. The conversational cloze test is a comprehensive tool to evaluate examinees' knowledge in pragmatics. One example of pragmatics from BCT is presented below:

(17)

Gisele: My dad's taking me to play the Dye-out game this Saturday. He asked if you would like to come.

Opal: \_ 22 \_.

Gisele: It's a popular game these days. The game is played by two teams and can be played inside or outside. You win the game by taking the other team's castle or “killing” all the players on the other team.

22. (A) But I'm not good at it  
(B) Your Dad has already invited me  
\*(C) But I've never heard of it before  
(D) I'd love to join you this Saturday

(From BCT 2013)

The clue of Question 22 is provided by the following sentences of Gisele's clear explanation about the Dye-out game. As the case shows, Opal elicits clarification by showing a statement of uncertainty. In this case, the cloze item is related to the knowledge of speech act in pragmatics; that is the statement functions as a requirement for further information about the Dye-out game in the conversation. However, there is no proper subtype for the usage of pragmatics; hence, a new test type of modification in *Category 4: Across Sentence, Within Text, 3-g: pragmatics* is added to the present study.

The sixth modification is made in Category 4. The cloze items of local context in the past mainly focus on the syntactic elements more than discourse structure; however, reading skills are included in the cloze section in recent years. The

test-takers need to find out the main idea across the sentences from the whole text.

One example from CAP is presented below:

(18)

Englishman Robert Scott is known for leading two trips to Antarctica. The first one made him a star; the second 38.

After his first successful trip, Scott decided to be the first person to stand on the South Pole. However, 39. He would be in a race with Roald Amundsen, from Norway. Both left their countries by ship in June of 1910 and arrived in Antarctica in January of 1911.

About ten months later, both teams started their trips down to the South Pole. 40. Amundsen's team used dogs, and Scott's team used horses. Because horses weren't good at traveling on snow, it took Scott's team 77 days to arrive at the South Pole. They got there on January 17, 1912, and were surprised to find that Amundsen was ahead of them. The news hit Scott very hard, but what he didn't know was that 41. His team began their long trip home with broken hearts. After days of terrible weather and little food, Scott lost his men one after another, and he himself was the last one to meet the end of his life. No one on his team lived to go back home and tell their story. It was only learned through Scott's diary.

38. (A) hurt his health

(B) opened his eyes

\*(C) cost him his life

(D) made his dream come true

(From CAP 2016)

In Example 18, to answer the Question 38, the test-takers need to understand the following description about the second trip to Antarctica, and then go back to find out the best option for its main idea. As the case stands, it's a big shift in cloze test. Test takers must acquire the ability of understanding how sentences are united together to build up a coherent text and identifying the main idea of the text. Therefore, a modification in *Category 4: 3-h discourse structure* is included in this study.

Table 3.3 illustrate the modified taxonomy of the present study.

Table 3. 3 Modified Taxonomy of the Current Study<sup>1</sup>

Clause-level Syntax (C1)	
1-S-a	determiner without text antecedent
1-S-b	conjunction below clause level
1-S-c	role-marking preposition
1-S-d	relative word in headless clause
1-S-e	complementizer
1-S-f	auxiliary
1-S-g	negation
1-S-h	copula Be & other intensive relationships
1-S-i	syntactic there or syntactic it + Be
1-S-j	syntactic elements in cleft or pseudo cleft
<b><i>1-L-a</i></b>	<b><i>referential determiner or deictic with clause</i></b>
<b><i>1-L-c</i></b>	<b><i>pronoun (coreferent with clause)</i></b>
<b><i>1-L-d</i></b>	<b><i>indefinite proform (including you as indefinite)</i></b>
<b><i>1-L-e</i></b>	<b><i>preposition other than role-marker</i></b>
Clause-level Lexis (C2)	
1-L-b	pre- & post determiners
1-L-f	collocation with clause
1-L-g	reiteration with clause
1-L-h	multipart item
1-L-i	qualifier in AdjP & AdvP
Across Clause, Within Sentence (C3)	
2-a	referential determiner or deictic across clauses
<b><i>2-b</i></b>	<b><i>coordinator</i></b>
2-c	pronoun (coreferent across clauses)
2-d	relative word in headed clause
2-e	collocation across clauses
2-f	reiteration across clauses
2-g	subordinator
2-h	<i>it</i> of extraposition
Across Sentence, Within Text (C4)	
3-a	referential determiner or deictic across sentences
3-b	conjunction between sentences
3-c	pronoun (coreferent across sentences)

<sup>1</sup> Mark the modifications by boldface and italic.

3-d	collocation across sentences
3-e	reiteration across sentences
3-f	textual discourse marker
3-g	<i>pragmatics</i>
3-h	<i>discourse structure</i>
<b><i>Extratextual - written passage and picture (ad, chart, map and others) (C5)</i></b>	
4-a	speech-role pronoun without textual antecedent
4-b	lexical item with little or <b>no clues from the text</b>

### Data Analysis

This section explains the data analysis framework of coding system and possible problems that may encounter during the data analysis. The procedure of the current study comprises the following steps, as illustrated in Figure 3.1.

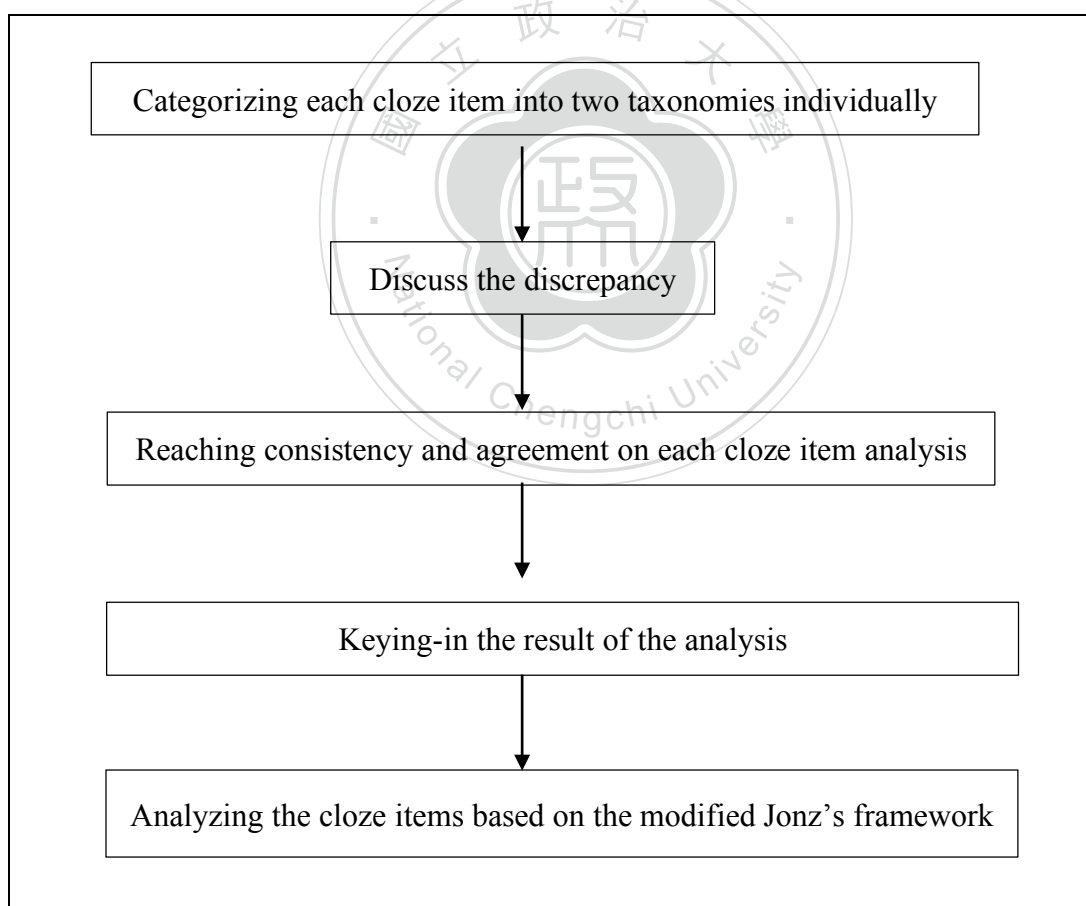


Figure 3. 1 Data Analysis Procedure of the Study

Firstly, to ensure the coding validity of the study, two coders categorized each cloze item according to Jonz's taxonomy: one is the researcher, the other is a senior high school English teacher, who is a full-time experienced senior high school English teacher, and has received MA linguistics training courses. Next, to establish the inter-rater reliability, the two coders needed to go through coding training to ensure both coders have consensus about the coding system. The training includes two phases. First, the researcher explained the purpose of the present study, and provided detailed explanation with sample examples of the coding system. Clear explanation and discussion of each category was further elaborated. Second, the researcher and the coder coded a small amount of the cloze items individually. The researcher and the coder then compared the results of each cloze item together during the coding practice to check if there was any discrepancy between the results. To solve the discrepancy, the two coders discussed the differences and resolved problems encountered. One example from CAP is presented below:

(19)

From : Lisa Clyne (lisaclyne@mail.com)

To : Mary Faber (maryfaber@mail.com)

Date : Thursday, April 2, 2015

Subject : Welcome

-----  
Dear Mary,

Uncle Billy and I are excited about your visit.

The weather has \_\_13.\_\_ these days. Like today, it was sunny in the morning, but rained heavily at noon. So we've \_\_14.\_\_ : If the weather is nice, I'm going to take you to Smith Farm. They're having a special horse show this week. I'm sure you'll like it. And \_\_15.\_\_. We can go shopping at Mimi's Department Store; there we can walk around without getting wet. Rain or shine, we hope you will have a good time here.

See you tomorrow at the train station at 9 : 00 a.m.

Love,

Aunt Lisa

13. (A) got better (B) been warm

\* (C) changed a lot (D) become wet and cold

14. \* (A) made two plans for you

(B) worried about your trip to here

- (C) decided to take you to a nice place
  - (D) prepared everything you've asked for
  - 15. (A) we'll agree with each other
  - \* (B) don't worry if it rains
  - (C) we all love animals very much
  - (D) don't forget to check the weather
- (From CAP 2015)

In Example 19, the researcher and the coder had an agreement about the type of category in Question 15 (Category 4), but different opinions about the subtype in this cloze item. The researcher classified the subtype of Question 15 into 3-b:

*conjunction between sentences*, and the reason was that the word “And” was utilized to connect the “two plans” mentioned in the text. On the other hand, the other coder classified Question 15 into 3-e: *reiteration across sentence*, and her reason was that the repetition of the word “rain” in the line 8 “Rain or shine” as a contextual clue. Then, the two coders discussed and reached the conclusion that the cloze item belonged to 3-b: *conjunction between sentences* in Question 15. The key to the final agreement was that “And” was used to connect the two plans, relating to the two conditional sentences “If the weather is nice” and “don't worry if it rains.”

The formal coding did not begin until a high degree agreement was reached in the results of the training. In the formal coding stage, the two coders followed the above procedure to examine each cloze item one by one. However, if any disagreement during the coding process could not be solved, the researcher must consult an expert to provide professional judgement.

Measurement of the extent to which data collectors (raters) agree upon the coding is called interrater reliability (McHugh, 2012). According to McHugh, the importance of rater reliability lies in the fact that it represents the extent to which the data collected in the study are analyzed in a consistent way. In the present study, the two coders categorized the target test items, and Cohen's Kappa coefficient was applied to measure the level of agreement between two coders.

The Kappa statistic varies from 0 to 1, and a value of 1 indicates perfect agreement and values less than 1 imply less than perfect agreement. The magnitude of Cohen's Kappa coefficient of the present study is .72. According to Viera and Garrett (2005), the Kappa value between .61 to .80 represents substantial agreement. This indicates that the consistency in the present study is moderately substantial. In addition, as explained in the framework of data analysis, the two raters discussed the discrepancy together until reached a concordance on each target items.

After the coding, the frequency of the five major categories and their subtypes was counted, and the percentage of all categories were shown in the next chapter. With the analysis of the frequency and percentage, the study presents what kind of language knowledge the cloze test constructors of BCT and CAP from 2010 to 2017 tend to measure, and the different language knowledge tested in both exams. Furthermore, the current study analyzes the passing rates of the cloze tests to get a more comprehensive understanding about test-takers' performances in such high-stakes exams. However, because the passing rates of BCT are not available according to RCPET, the present study would only analyze the passing rates of CAP since 2014.



## CHAPTER FOUR

### RESULTS AND DISCUSSION

In this chapter, the results of the analysis are presented and discussed in response to the two research questions. The first section presents the data and discusses the result of the distribution of Jonz's five categories in BCT and CAP. The second section is to present and discuss the result of the distribution of Jonz's 37 subtypes in BCT and CAP. The first and the second sections aim to answer the first research question. The last section compares the distribution of five categories and 37 subtypes of BCT and CAP to show the differences and similarities in order to answer the second research question.

#### *Results of the Analysis on the Five Categories*

Jonz's five categories includes: (1) Clause-Syntax, (2) Clause- Lexis, (3) Across Clause, Within Sentence, (4) Across Sentence, Within Text, and (5) Extratextual representing the scope that a test item covers in the text. The outcome and result of Jonz's 5 categories are discussed in this section. The results present the distribution of cloze items of the senior high school entrance examination from 2010 to 2017, including BCT and CAP, under Jonz's 5 categories. The total exam texts are 20, and cloze items are 66 in both exams during the past 8 years.

The result of the distribution of five categories is presented in Table 4.1.

Table 4. 1 The Distribution of Items in the Five Categories

Category	Number	Percentage
C1: Clause-Syntax	3	4.5%
C2: Clause- Lexis	1	1.5%
C3: Across Clause, Within Sentence	13	19.6%
C4: Across Sentences, Within Text	47	71.2%
C5: Extratextual	2	3%
<b>Total</b>	<b>66</b>	<b>100%</b>

From this table, we can observe the cloze items of the senior high school entrance examination primarily focused on *C4-Across Sentence, Within Text*. Forty-seven (71.2%) out of 66 tested cloze items are categorized into C4. This category aims to examine whether candidates are capable of choose the correct answers from the contextual clues across the sentences. One example of C4 is presented below:

(20)

Do you eat meat? Well, if you do, then you might find our next news interesting. A study says that the world has a 33. need for meat. In 1960, the world ate 64 million tons of meat, about 21 kg for each person. In 2007, the number rose to 268 million tones, about 40 kg for each person. At the same time, 34. In the 1960s, beef was high on the menu. Of the meat that was eaten, 40% was beef. In 2007, pork became the star. Poultry also became popular, going up from 12% to 32%, thanks to people's worries about their health these years. And 35.? You're guessing the U.S.A., right? The answer is Luxembourg! In 2007, each Luxembourger ate about 137 kg of meat! Second to Luxembourgers are Americans. In 2007, each American ate about 126 kg! Now, enough with the numbers. I'm playing you a song called *Currywurst*. The singer sings about his love for the meat dish of the same name. Enjoy!

33. (A) falling (B) special \*(C) growing (D) common

34. (A) we have changed our way of cooking meat

(B) new kinds of meat have come on the market

\*(C) there have been changes in the list of favorite meats

(D) doctors have been worried that we eat too much meat

35. (A) where does the world's best meat come from

(B) which country is the world's biggest meat maker

(C) where can you eat the world's most delicious meat

\*(D) which country uses up the most meat for each person

(From CAP 2014)

In Example 20, questions 33, 34 and 35 were categorized into C4. To choose the correct answer for the three questions, candidates were required to look for more information in the following descriptions. In Question 33, the missing word is related to the main idea about what the study suggests. To give the correct answer, the test takers have to read the following two sentences about the amount of meat consumed, and then to get the contextual clue of a key word “*rose*” in the fifth sentence, which means the need for the meat is increasing. Therefore, Answer C “*growing*” is the correct answer.

Similarly, the missing clause is a general statement functioning as a topic sentence in Question 34. The correct answer is a statement that connects the next four supporting sentences to construct a logical text. Therefore, the candidates must recognize the following details in the supporting sentences for the idea of the changes of favorite meat. The following sentences mentioned some contextual clues related to the idea of favorites meat, which are “*high on the menu, became a star, popular, going up from 12% to 32%.*” With the contextual clues, the sentence of Option C is the most logical and coherent one.

In the same way, the options for Question 35 are four questions. The test takers need to read the following answers given in the next sentences that provide the contextual clues for the question asked. The answers in the following sentences mentioned two countries, Luxembourg and America, and how much meat the inhabitants ate in these countries in 2007. In this way, the test takers could remove the irrelevant questions, and thus identify the appropriate question is about a country where the most meat was consumed by each person (Answer D).

In sum, the reader can find the textual relationships (C4) are involved in the above example. In other words, the test constructors put strong emphasis on the measurement of discourse structure. The test items in discourse structure include main idea, and topic

sentence; both are the important elements in constructing a well-organized textual relationship of a coherent text.

The second highly involved category is *C3: Across Clause, Within Sentence*.

Thirteen (19.6%) out of 66 tested cloze items were categorized into C3. One example of C3 is presented below.

(21)

Long long time ago, a girl named Doris lost her ring. She felt so bad that she could not sleep well. When Doris turned over in her bed, she found the black horse in the picture on the wall 21.. She was very surprised, so she got out of bed to have a closer look. When Doris touched the head of the horse, she 22. into the picture! Now Doris was sitting on the back of the running horse! It was so exciting to be riding a horse in the beautiful sky, but Doris did not know where the horse 23. her. Finally, they stopped in front of an old church. There Doris found a gift bag. She opened the bag and saw her ring inside. Doris was very happy. When she put on the ring, she woke up from her dream. At that moment, Doris could not believe her eyes – the ring 24. there back on her finger!

21.(A) is moving (B) has moved \*(C) was moving (D) will move

22.(A) is pulled \*(B) was pulled (C) has been pulled

(D) would be pulled

23.(A) took (B) is taking (C) has taken \*(D) would take

24.(A) is \*(B) was (C) has been (D) would be

(From BCT 2011-1)

In Example 21, questions 21 to 24 were categorized into C3, that is, the contextual clues of these four questions were across clauses but within a sentence. The target items of questions 21 to 24 tested on tense forms. The candidates were required to have a correct time frame to the story. In Example 21, the setting of the text is about what happened to a girl named Doris long time ago. The test takers needed to identify the proper verb tense forms from the story. In Question 21, the test takers were asked to identify the tense form of the missing verb “*move*.” The word is in a sentence which connected by two clauses; one is a main clause, and the other is a subordinate clause. The subordinating conjunction “*when*” indicated the relationship about the time (Doris turned over in her bed) and action (what she found). Therefore, the test takers could find the most appropriate tense is Answer C – past progressive.

Similarly, in Question 22, the test takers were asked to identify the tense form of the missing verb “*pull*.” The subordinating conjunction “*when*” in Question 22 indicated the relationship about the time (Doris touched the head of the horse) and action (what happened). The test takers thus could identify the correct tense form is Answer B – passive voice in past tense.

Besides, the test takers could follow the contextual clues to get the correct tense forms within a sentence in questions 23 and 24 as well. In Question 23, the conjunction “*but*” connected two clauses in past tense. Then, options B and C could be deleted. Afterwards, according to the sentence “*but Doris did not know where the horse 23 her.*” The appropriate answer “*would take*” in Answer D is better than “*took*” in Answer A. In the same case, the test takers could identify the time frame in question 24 by the contextual clues within a sentence, including “*At that moment*” and “*Dorise could not believe...*” to choose the correct tense form is “*was*” in Answer B.

Referring to Table 4.1 about the coverage of the five categories, the reader can find C3 and C4 together account for 60 items (90.8%) of all, and the percentage is rather higher than the other three categories (C1, C2, and C5). The high proportion indicates that the primary language knowledge tested in BCT and CAP is about the higher-order global knowledge, which means the candidates must be able to deal with the missing target forms by integrating the information across clauses or sentences to complete a coherent and united text.

Table 4.1 also shows that only one item out of 66 was categorized into *C2-Clause Lexis*; two out of 66 was categorized into *C5- Extratextual*; three out of 66 was categorized into *C1-Clause syntax*. The results suggest that only 6 (9%) out of 66 cloze items focused on low-order local knowledge in BCT and CAP among the past 8

years. It is apparent that the cloze items for local understanding are getting less in senior entrance examination in Taiwan.

*The Results of Jonz's 37 Subtypes in BCT and CAP*

This section presents the result of Jonz's 37 subtypes in the senior high school entrance examination in the past 8 years. Firstly, Table 4.2 displays the numbers and percentages of the matched items.



Table 4. 2 The Distribution of Jonz's Subtypes in BCT and CAP

Clause-level Syntax (C1)		Total Numbers and Percentages
1-S-a	determiner without text antecedent	0 (0%)
1-S-b	conjunction below clause level	0 (0%)
1-S-c	role-marking preposition	0 (0%)
1-S-d	relative word in headless clause	0 (0%)
1-S-e	complementizer	0 (0%)
1-S-f	auxiliary	0 (0%)
1-S-g	negation	0 (0%)
1-S-h	copula Be & other intensive relationships	0 (0%)
1-S-i	syntactic there or syntactic it + Be	0 (0%)
1-S-j	syntactic elements in cleft or pseudo cleft	0 (0%)
1-L-a	referential determiner or deictic with clause	3 (4.5%)
1-L-c	pronoun (coreferent with clause)	0 (0%)
1-L-d	indefinite proform (including you as indefinite)	0 (0%)
1-L-e	preposition other than role-marker	0 (0%)
Clause-level Lexis (C2)		
1-L-b	pre- & post determiners	0 (0%)
1-L-f	collocation with clause	1 (1.5%)
1-L-g	reiteration with clause	0 (0%)
1-L-h	multipart item	0 (0%)
1-L-i	qualifier in AdjP & AdvP	0 (0%)
Across Clause, Within Sentence (C3)		
<b>2-a</b>	<b>referential determiner or deictic across clauses</b>	<b>12 (18.1%)</b>
2-b	coordinator	0 (0%)
2-c	pronoun (coreferent across clauses)	1 (1.5%)
2-d	relative word in headed clause	0 (0%)
2-e	collocation across clauses	0 (0%)
2-f	reiteration across clauses	0 (0%)
2-g	subordinator	0 (0%)
2-h	it of extraposition	0 (0%)
Across Sentence, Within Text (C4)		
<b>3-a</b>	<b>referential determiner or deictic across sentences</b>	<b>14 (21.2%)</b>
3-b	conjunction between sentences	2 (3%)
3-c	pronoun (coreferent across sentences)	1 (1.5%)
3-d	collocation across sentences	2 (3%)

3-e	reiteration across sentences	4 (6%)
3-f	textual discourse marker	3 (4.5%)
3-g	pragmatics	2 (3%)
<b>3-h</b>	<b>discourse structure</b>	<b>19 (28.7%)</b>
Extratextual - written passage and pictures (C5)		
4-a	speech-role pronoun without textual antecedent	0 (0%)
4-b	lexical item with little or no clues from the text	2 (3%)
Total		66 (100%)

From the above table, the most frequently-tested subtype of all the cloze items is 3-h: *discourse structure*, nineteen cloze items (28.7%) out of 66 were classified into this target item. As discussed in Example 20, another example of 3-h: *discourse structure* is presented below:

(22)

### NEWS HUNT

Kieran Hardy 11/07/2013

For years, we thought our Earth was the only blue dot up there. Now another has been found. Its name is HD189773b. HD189773b is an exoplanet, a planet outside our Solar System, and is one of the nearest exoplanets to Earth. Even so, 16.. Here's why: It is 63 light years away. That means it is 370,440,000,000,000 miles from us. Even if we fly at 3,500 miles an hour, it will take more than 12 million years to get there.

17.: It is much bigger, it is made of gas and it is burning hot. In heat as great as 1,000°C, life is not possible. What is worse, it rains glass. If 1,000°C does not kill you, glass rain will.

Even so, finding this blue giant 18.. It is the first time that we have been able to see the color of an exoplanet. The color of a planet gives us ideas about what is happening on it. While we have a long way to go before we find a new planet to live on, finding a blue dot is a good start.

16. \*(A) we cannot really call it a neighbor

(B) we do not know anything about it yet

(C) we might not be able to stay there for long

(D) we are not sure how long it takes to get there

17. (A) What's more, it is not water that makes HD189773b look blue

(B) With its blue color, HD189773b could be a second Earth for us

\*(C) Except for its blue color, HD189773b is nothing like our home planet

(D) Because of its blue color, people guess there might be life on HD189773b

18. \*(A) gives us hope

(B) took hard work

(C) has changed our life (D) helps us know more about Earth

(From CAP 2015)



As Example 22 shows, to get the correct answers to questions 16 to 18 are either topic sentences or main idea that govern the following descriptions. In question 16, the next supporting details are about the explanation of how far the exoplanet HD189773b from Earth and how much time it takes us to fly to get there in the first paragraph. With the explanation, it gave important information that although HD189773b is one of the nearest exoplanet to Earth, it is still has a very long distance from our Earth. Therefore, the appropriate statement to connect the following details is Option A.

For Question 17, the supporting sentences in the second paragraph introduce some information of exoplanet HD189773b: the component, the temperature, and the special phenomenon about raining glass. From these contextual clues, the test takers could tell that HD189773b is nothing like Earth except for its blue color. Consequently, the test takers could remove the irrelevant sentences in Option A which talked about the color and the water, and Option B and Option D which were about the idea of resemblance between Earth and HD189773b. Thus, the most appropriate topic sentence is Answer C.

Similar to Question 17, the cloze item in Question 18 is about the main idea in the third paragraph. The following details provide the contextual clues about “*a new planet to live on*” and “*a good start*”. Therefore, Option A “*give us hope*” is the main idea in the topic sentence of the third paragraph.

As the case stands in Example 22, it requires the test takers to pay more attention to the subtle textual relationship among the sentences, which form the discourse structure to make the text logical and coherent.

The second frequently-tested subtype is 3-a: *referential determiner or deictic across sentences*, fourteen items (21.2%) out of 66 were classified into this target item. 3-a is utilized to measure the use of tense forms in English. One example of deictic across sentences is presented below:

(23)

People in Quiet Town are very sad that their old friend, Time River, is facing the end of its life next October. The river 17. a lot to the people in the town. They all grow up and live with it all their lives.

17. \*(A) means (B) is meaning (C) will mean (D) meant

(From BCT 2010-2)

In Example 23, the test takers should identify the correct tense for the missing word.

The clue for Question 17 lies between the sentences that describes the situation that the river faces “now” and the significance of the river to the people in the town. All the sentences in the same time frame are in the present tense. With these descriptions, the test takers could identify the best tense for the target item was present tense.

Besides, because the subject was a single noun “*the river*”, the verb “*mean*” must be changed into “*means*” according to the grammar of present tense.

Another example of referential determiner across sentences is presented below:

(24)

Mr. Hutman owns a restaurant. One day, one of his two cooks ran away with some money. Mr. Hutman was sad and worried because the next day was Saturday, and the restaurant 19. very busy then. So that night he called his friends, but could not find anyone to help.

On Saturday, many people came to the restaurant for lunch. 20. cook got so busy that he wanted to leave, too. 21., one of the waiters, Henry, told Mr. Hutman that he could help in the kitchen. People who had lunch in the restaurant that day enjoyed the food Henry cooked. Mr. Hutman was very happy and made Henry a new cook of the restaurant.

19. \*(A) would be (B) is (C) has been (D) will be

20. (A) One (B) Each (C) Another \*(D) The only

21. (A) Also (B) In fact \*(C) Luckily (D) For example

(From BCT 2010-1)

In Example 24, the target tested word of Question 20 is the referential determiner.

The candidates could get the contextual clue by referring to the second sentence in the first paragraph “*One day, one of his two cooks ran away with some money.*” Since one cook ran away, the second cook is the only one in the restaurant (Answer D).

Overall, the second frequently-tested subtypes 2-a and 3-a primarily focus on the tense forms. Twenty-five items out of 66 tested in tense forms, and only one item

tested in referential determiner. However, the learning of referential determiners plays a crucial role in the four skills in English (listening, speaking, reading, and writing).

Referring back to Table 4.2, the reader can observe that there are totally 37 subtypes in Jonz's taxonomy, but only 13 (35.1%) out of the 37 subtypes were tested in BCT and CAP, that is, 24 (64.9%) out of 37 were absent from both exams. Moreover, compared to the subtypes in C1, C2, C3 and C5, C4 has as many as 47 cloze items classified into its 8 subtypes, that is, every subtype of C4 is included in both exams in the past 8 years. In short, the senior high school entrance examinations in the past 8 years do not incorporate a balanced range of language subtypes in cloze test section, and the test developers may distribute the other important language test items in other sections, especially in the section of single-questions in the first part of the exam.

#### *The Distribution of Jonz's 5 Categories in Each Year of BCT and CAP*

The purpose of this section is to discuss the distribution of Jonz's 5 categories among the past 8 year. The researcher present the results to show the changes of the measurement of language knowledge in cloze tests in local context. Table 4.3 displays the result of the distribution of Jonz's 5 categories of each year.

Table 4. 3 The Distribution of Jonz's Five categories from 2010 to 2017

	Category One	Category Two	Category Three	Category Four	Category Five	Total
2010-1BCT	0 (0%)	0 (0%)	<b>1(16.6%)</b>	<b>5(83.3%)</b>	0 (0%)	6(100%)
2010-2BCT	<b>1(14.2%)</b>	<b>1(14.2%)</b>	0 (0%)	<b>5(71.4%)</b>	0 (0%)	7(100%)
2011-1BCT	0 (0%)	0 (0%)	<b>4(66.6%)</b>	0 (0%)	<b>2(33.3%)</b>	6(100%)
2011-2BCT	<b>2(28.5%)</b>	0 (0%)	<b>3(42.8%)</b>	<b>2(28.5)</b>	0 (0%)	7(100%)
2012 BCT	0 (0%)	0 (0%)	<b>1(16.6%)</b>	<b>5(83.3%)</b>	0 (0%)	6(100%)
2013 BCT	0 (0%)	0 (0%)	<b>2(33.3%)</b>	<b>4(66.6%)</b>	0 (0%)	6(100%)
2014 CAP	0 (0%)	0 (0%)	0 (0%)	<b>7(100%)</b>	0 (0%)	7(100%)
2015 CAP	0 (0%)	0 (0%)	0 (0%)	<b>6(100%)</b>	0 (0%)	6(100%)
2016 CAP	0 (0%)	0 (0%)	<b>2(28.4%)</b>	<b>5(71.4%)</b>	0 (0%)	7(100%)
2017 CAP	0 (0%)	0 (0%)	0 (0%)	<b>8(100%)</b>	0 (0%)	8(100%)
Total	3	1	13	47	2	66

As shown in Table 4.3, there are 6 exams and 38 cloze items in BCT from 2010 to 2013. Among them, twenty-one (55.2 %) items belong to C4, eleven (39.2 %) items are at C3, three (7.8 %) items are at C1, two (5.2%) items are at C5 and only one (2.6 %) item at C2. From the above result of analysis, the highest coverage among the five categories is C4.

Besides, BCT items covered from C1 to C5. As shown in Table 4.3, most items categorized into C4, especially in 2010-1 BCT and 2012 BCT, each year account for 83.3%. The reader can find that C4 was absent only in 2011-1 BCT. More items are categorized in C3 in 2011-1 BCT. Besides, BCT items are distributed into three categories in 2010-2 BCT and 2011-2 BCT. When comes to 2012 to 2013, the items categorized into C4 started to increase. With the above analysis, it is apparently that BCT diversified more measurement of language ability by covering different language knowledge.

As for CAP from 2014 to 2017, there are 4 exams and 28 cloze items in

Table 4.3. We can see up to 26 (92.8 %) items are in C4, only 2 (7.1%) items are in C3, and no items are in C1, C2 and C5. C4 accounts for the largest proportion among the five categories. What's more, it is noteworthy that in the years of 2014, 2015 and 2017, all cloze items belong to C4. The diversification of cloze items among the five categories is unbalanced in CAP. The reason can be that certain subtypes of grammatical structures are tested by other sections of the test.

According to the website of the official item analysis reported by Research Center for Psychological and Educational Testing, CAP is more difficult than BCT (<https://cap.nace.edu.tw>). For BCT, there are more items falling on the relatively easy level, and items of CAP are in moderate of difficulty, which means the proportion in the degree of difficulty is equally distributed among the test in CAP. What's more, the length of the cloze texts and readability also reflect the shift of the degree of difficulty in BCT and CAP. Table 4.4 and Table 4.5 display the numbers of words and readability in both exams.

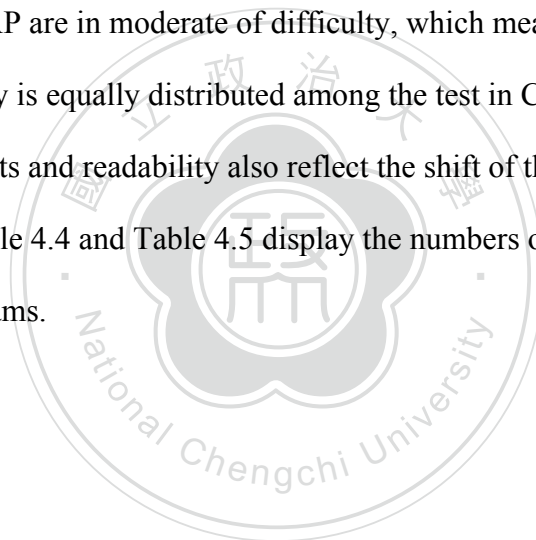


Table 4. 4 The Numbers of Words and Readability in BCT

BCT (Year/Question Number)	Words	Readability (Flesh-Reading Ease)
2010-1 (19-21)	118	77.1
(22-24)	142	91.1
2010-2 (17-20)	167	85
(21-23)	157	73.2
2011-1 (21-24)	163	90.2
(25-26)	109	85.8
2011-2 (21-24)	118	90.7
(25-27)	133	90.2
2012 (21-23)	88	90.8
(24-26)	119	77.9
2013 (19-21)	99	68
(22-24)	150	75.1
	<b>Average: 130.2</b>	<b>Readability Index: 82.9</b>

Table 4. 5 The Numbers of Words and Readability in CAP

CAP (Year/Question Number)	Words	Readability (Flesh-Reading Ease)
2014 (33-35)	193	71.6
(57-60)	169	89.1
2015 (13-15)	84	85.3
(16-18)	210	62.7
2016 (35-38)	147	78.3
(39-41)	133	64
2017 (34-37)	224	72.6
(38-41)	214	80
<b>Average</b>	<b>Average: 171.7</b>	<b>Readability Index: 75.4</b>

As shown in Table 4.4 and Table 4.5, the average length of the text is 130.2 words in BCT, and 171.7 words in CAP. We can conclude that most of the texts in CAP are longer than BCT. In terms of readability, BCT has a readability index about 82.9. Five texts out of 12 have a general readability score of high 90s, and the highest (easiest) one is 91.1. The readability scores of two texts were between 85 to 85.8; four between

73.2 to 73.9. One text was in the lowest (the most difficult) score of 68 in the text of Questions 19-21 in 2013.

The readability index of CAP is about 75.4. Three texts out of 8 have a general readability score between 80 to 89.1, and the highest (easiest) one is 89.1 in 2014. The readability scores of 3 texts out of the 8 are between 71.6 to 78.3; two were in the scores between 62.7 to 64. The lowest (the most difficult) score is 62 in the text of Questions 16-18 in 2015. From the analysis results, the length and the readability indicate that CAP is more difficult than BCT. Once the exam enhances the degree of difficulty, the test takers are facing a more challenging exam than before because they are required to deal with more complex texts.

*Results of the Distribution of Jonz's 37 Subtypes in BCT and CAP*

This section discusses the distribution of the 37 subtypes. Table 4.6 presents the distribution of BCT and CAP items.

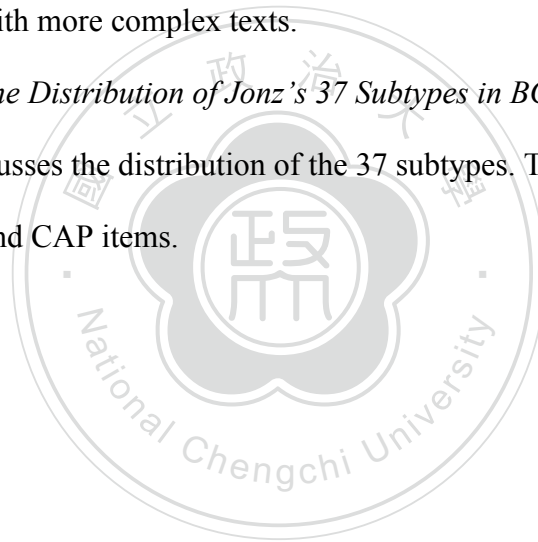


Table 4. 6 The Distribution of Jonz's 37 Subtypes in BCT and CAP

Clause-level Syntax (C1)		Total Numbers and Percentages	
		BCT	CAP
1-S-a	determiner without text antecedent	0 (0%)	0 (0%)
1-S-b	conjunction below clause level	0 (0%)	0 (0%)
1-S-c	role-marking preposition	0 (0%)	0 (0%)
1-S-d	relative word in headless clause	0 (0%)	0 (0%)
1-S-e	complementizer	0 (0%)	0 (0%)
1-S-f	auxiliary	0 (0%)	0 (0%)
1-S-g	negation	0 (0%)	0 (0%)
1-S-h	copula Be & other intensive relationships	0 (0%)	0 (0%)
1-S-i	syntactic there or syntactic it + Be	0 (0%)	0 (0%)
1-S-j	syntactic elements in cleft or pseudo cleft	0 (0%)	0 (0%)
1-L-a	referential determiner or deictic with clause	<b>3 (7.8%)</b>	0 (0%)
1-L-c	pronoun (coreferent with clause)	0 (0%)	0 (0%)
1-L-d	indefinite proform (including you as indefinite)	0 (0%)	0 (0%)
1-L-e	preposition other than role-marker	0 (0%)	0 (0%)
Clause-level Lexis (C2)			
1-L-b	pre- & post determiners	0 (0%)	0 (0%)
1-L-f	collocation with clause	<b>1 (2.6%)</b>	0 (0%)
1-L-g	reiteration with clause	0 (0%)	0 (0%)
1-L-h	multipart item	0 (0%)	0 (0%)
1-L-i	qualifier in AdjP & AdvP	0 (0%)	0 (0%)
Across Clause, Within Sentence (C3)			
2-a	referential determiner or deictic across clauses	<b>10 (26.3%)</b>	<b>2 (7.1%)</b>
2-b	coordinator	0 (0%)	0 (0%)
2-c	pronoun (coreferent across clauses)	<b>1 (2.6%)</b>	0 (0%)
2-d	relative word in headed clause	0 (0%)	0 (0%)
2-e	collocation across clauses	0 (0%)	0 (0%)
2-f	reiteration across clauses	0 (0%)	0 (0%)
2-g	subordinator	0 (0%)	0 (0%)
2-h	it of extraposition	0 (0%)	0 (0%)
Across Sentence, Within Text (C4)			
3-a	referential determiner or deictic across sentences	<b>6 (15.7%)</b>	<b>8 (28.5%)</b>
3-b	conjunction between sentences	0 (0%)	<b>2 (7.1%)</b>
3-c	pronoun (coreferent across sentences)	<b>1 (2.6%)</b>	0 (0%)



3-d	collocation across sentences	2 (5.2%)	0 (0%)
3-e	reiteration across sentences	3 (7.8%)	1 (3.5%)
3-f	textual discourse marker	3 (7.8%)	0 (0%)
3-g	pragmatics	2 (5.2%)	0 (0%)
3-h	discourse structure	4 (10.5%)	15 (53.5%)
Extratextual - written passage and picture (C5)			
4-a	speech-role pronoun without textual antecedent	0 (0%)	0 (0%)
4-b	lexical item with little or no clues from the text	2 (5.2%)	0 (0%)
<b>Total</b>		<b>38 (100%)</b>	<b>28 (100%)</b>

Firstly, only BCT has C1 and C2 clause level items. From Table 4.6, three items and one item are categorized into C1 and C2 in BCT. They totally account for 10.4 %. As for CAP, no items are categorized into C1 and C2. It is apparent that the focus of the tested points within a clause level knowledge in CAP is completely absent. Second, both exams have C3 items, but BCT are far more than CAP. The subtypes belong to C3 means that the correct answer can be determined from the contextual clue within the same sentence. As shown in Table 4.6, the most frequently-measured subtype in BCT is *2-a referential determiner or deictic across clauses* in C3. Ten items (26.3%) out of 38 target at the knowledge of tense forms. On the contrary, only two items (7.1%) out of 28 in CAP belong to the Subtype 2-a in C3. The numbers and percentages of C3 items in CAP are less than BCT.

Third, both exams have items in C4 but with different distribution. Six (15.7%) items out 38 belong to Subtype 3-a in BCT, and eight (28.5%) out of 28 are classified into Subtype 3-a in CAP. The cloze items in 3-a focus on the syntactic knowledge across sentences, including tense forms and referential determiner. The numbers of 3-a are similar in both exams, and the percentage in CAP is higher than BCT.

Furthermore, as shown in table 4.6, there are 28 cloze items in total in CAP during 2014 to 2017. Among them, the numbers and percentages of the most frequently-measured subtype is 15 items (53.5%) of *3-h: discourse structure* in C4.

As for BCT, the number and percentage of Subtype 3-h is 4 (10.5%) items out of 38. The proportion of the measurement of language knowledge in discourse structure in CAP is much higher than BCT.

Lastly, C5 involves contextual clues usually present the texts together with pictures (ads, charts and others). The result of the distribution presents 2 (5.2%) items out of 38 are categorized into C5 in BCT, and no items in CAP belong to this category. The BCT questions included dialogues, ads, and letters demonstrated a wide range of situations in daily life. However, those genres are not employed in the cloze test section of CAP. CAP tend to use long written articles.

The passing rates of the cloze section can provide further information about the difficulty level of both exams. However, because the passing rates of BCT are not available from the website of RCPET, the present study thus only analyzes the data in CAP. Table 4.7 presents the data and distribution of the passing rates of the cloze section in CAP which announced by RCEPT from 2014 to 2017 (Appendix D).

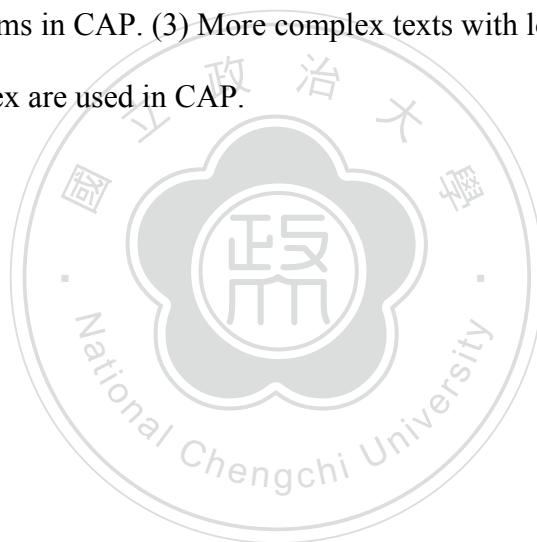
Table 4. 7 The Distribution of the Passing Rates of the Items in CAP

Passing Rates	Numbers and Percentages
0.63	2 (7.1%)
<b>0.5-0.59</b>	<b>13 (46.4%)</b>
<b>0.4-0.49</b>	<b>10 (35.7%)</b>
0.32-0.38	3 (10.7%)
<b>Total items</b>	<b>28 (100%)</b>

As Table 4.7 displays, only 2 (7.1%) out 28 have the highest passing rate of 0.63, and the number of the items with the passing rates between 0.5 and 0.59 is 13 (46.4%) out of 28; ten (35.7%) is between 0.4 to 0.49; three (10.7%) is between 0.32 to 0.38. The results of the distribution indicate that the items with the passing rates between 0.59 to

0.49 reach up to 82.1 % in total in CAP from 2014 to 2017. It shows the degree of difficulty among the four years is between medium easy to medium difficult<sup>2</sup>. Also, three items even have the passing rates below 0.4. The above analysis of the passing rates in CAP shows that the cloze test may be a quite challenging section for some test-takers.

In sum, the results of analysis of Jonz's 5 categories and 37 subtypes indicate that there are great differences between BCT and CAP in cloze test section in three ways: (1) The measurement of language knowledge changes from clause level in BCT to global comprehending across sentences in CAP. (2) The increase of remarkable discourse structure items in CAP. (3) More complex texts with longer length and harder readability index are used in CAP.



---

<sup>2</sup> According to CEEC, the following is the interpretation of passing rates: 0-0.25 Difficult, 0.26-0.50 Medium Difficult, 0.51-0.75 Medium Easy, 0.76-1.00 Easy.



## **CHAPTER FIVE**

### **CONCLUSION**

This chapter presents a conclusion to wrap up the entire study based on the results and discussion of the current study. The first section presents a summary of the major findings to answer the two research questions. The second section provides teaching implications to language teachers in junior high school. Lastly, the limitations of the present study are discussed.

#### **Summary of the Major Findings**

Washback effect is “the extent to which the introduction and use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning” (Messick, 1996, p. 241). The washback effect can be negative in what happens in the test, and it also can be positive if the test is well-designed. Therefore, the test result of both high-stakes senior high school entrance examinations is of vital importance which decides the future school for study to each 9<sup>th</sup> graduate. The present study is thus motivated to analyze the cloze test section to find out what the language knowledge is measured in both exams in the past 8 years. The reason to choose the cloze test section is that it targets at integrated comprehension skills, and it is used widely in many testing contexts in Taiwan.

The materials in the present study are collected from BCT and CAP in the past 8 years. The researcher applies the taxonomy of Jonz’s study in 1990 to analyze the features of the cloze items in both exams. The results of the present study can be summarized through the two research questions as follows.

### **RQ1. What kind of language is measured by cloze items of BCT and CAP from 2010 to 2017?**

Firstly, the present study analyzed the cloze items in the past 8 years of both exams by classifying them into 5 categories: (1) Clause-Syntax, (2) Clause-Lexis, (3) Across Clause, Within Sentence, (4) Across Sentence, Within Text, and (5) Extratextual. The finding is that both exams have their cloze tests focus on C3 and C4. The high proportion of C3 and C4 items indicate that the primary language knowledge tested is more about the higher-order knowledge which means the candidates must be able to deal with the target items by integrating the contextual clues across sentential level. The tested items about restricted points in local understanding are getting less in the past 8 years.

Secondly, the subtype of Jonz's taxonomy 3-*h*: *discourse structure* of C4 has increased greatly in the past 8 years. It requires the test takers to pay more attention on the textual relationship to complete a logical and coherent text. The focus of discourse structure also corresponds to the result that both exams put more emphasis on the items beyond sentences.

Lastly, the second frequently-tested subtype is 3-a. It is utilized to measure the language knowledge about tense. The preference of the syntactic knowledge in tense of both exams reflects that in the present language learning context, the syntactic knowledge in tense take a large proportion in English grammar teaching at the stage of junior high school.

### **RQ2. What are the differences between cloze items of BCT and CAP from 2010 to 2017?**

After analyzing each cloze item in BCT and CAP, three findings are obtained from the analysis results in the previous chapter. First, the cloze section of BCT covers categories of C1, C2, C3, C4 and C5. CAP has a strong preference in C4, and

no items belong to C1, C2 and C5. The absent but important language ability may be distributed in other sections. Secondly, the most frequently-tested subtype of BCT is 2-a in C3 (p. 54), and it focus on the syntactic knowledge in tense. For CAP, the most frequent-measured subtype is 3-h, and it requires the candidates to be familiar with the textual relationship in discourse. As the analysis results present, it is a big shift in cloze items between BCT and CAP; that is, from syntactic knowledge to discourse structure.

Thirdly, according to RCEPT, CAP is more difficult than BCT. The present study finds the same conclusion as announced by RCEPT by analyzing the length and the readability of the texts in both exams. The average length of the text in CAP is longer than BCT (171.7 words vs. 130.2 words). Also, the readability index in CAP are lower than BCT (75.4 vs. 82.9). Besides, the passing rates of the cloze section in CAP from 2014 to 2017 mainly fall in the range between 0.4 to 0.59, account for 82.1%. We can thus conclude that the cloze section may be a challenging part for the test takers.

### **Suggestions and Implications**

The tendency of CAP in testing the knowledge of discourse structure actually echoes the focus of new Grade 1-12 Curriculum Guidelines. With the changes of the focus in the test, the adjustment and modification in English curricula at the stage of junior high school should be made accordingly. From the analysis results, three suggestions are provided to English teachers in junior high school: (1) The focus of English teaching activity should shift from the explanation of vocabulary and grammar patterns to the development of the ability of textual knowledge. (2) The teaching materials should include more on comprehensible extensive reading. It is of vital importance to develop learners' reading ability in integrating the longer and complex information. (3) The change of the cloze items from syntactic knowledge to

discourse structure in CAP is a big challenge for students as well as teachers. The English teachers should engage in further study in the discourse field in the future. In this case, teacher can face the requirement in the teaching job and provide due assistance for students in the English language learning.

### **Limitations**

The first limitation lies in the lack of the feedbacks from the test takers. The present study thus is not comprehensive enough to present the impact of both exams. Future studies may elicit the feedback and perception from both language teachers and students as test users through questionnaire or interview to acquire more information.

The second limitation is in the lack of the data of passing rates in BCT. The analysis results of the present study show the difficulty level of CAP by interpreting the passing rates announced by RCEPT. However, the passing rates of BCT are not available to compare them with CAP's. Without the comparison, the study cannot show the actual differences of test takers' performance in the cloze section of both exams.



## References

- Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly*, 13(2), 219-227.
- Alderson, J. C. (1980). Native and nonnative speaker performance on cloze tests. *Language Learning*, 20, 59-76.
- Alderson, J.C. (2005). *Assessing reading*. (5<sup>th</sup> ed). Cambridge: Cambridge University Press.
- Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly*, 16(1), 61-70.
- Bachman, L. F. (1985). Performance on cloze with fixed-ratio and rational deletions. *TESOL Quarterly*, 19(3), 535-556.
- Bailey, K. M. (1998). *Learning about language assessment: Dilemmas, decisions, and directions*. Boston: Heinle & Heinle.
- Bensoussan, M. (1990). Redundancy and the cohesion cloze. *Journal of Research in Reading*, 13, 18-17.
- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy*. (2nd ed.). NY: Longman.
- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program design*. Boston: Heinle & Heinle.
- Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. NY: McGraw-Hill.
- Chang, W. C. (1994). A brief analysis of the principles for constructing cloze items. *Newsletter for Teaching the Humanities and Social Studies*, 5(4), 69-80.
- Chavez-Oller, M. A., Chihara, T., Weaver, K. A., & Oller, J. W. (1994). When are cloze items sensitive to constraints across sentences? *Language Learning*, 35(2), 181-206.
- Chou, S. Y. (2009). *A Study of Cloze Test Items in Scholastic Aptitude English Test and Department Required English Test*. Unpublished master's thesis, National Chung Cheng University, Taipei, Taiwan.
- Chen, W. Y. (2008). *The relationship between the rational cloze test and the discourse structure test*. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Cheng, H. Y. (2007). *A study of the cohesion items in the cloze tests of SAT and AST*. Unpublished master's thesis, National Chengchi University, Taipei, Taiwan.
- Dale, E. & Chall, J. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27, 11-20.
- Dastjerdi, H. V., & Talebinezhad, M. R. (2006). Chain-processing deletion procedure in cloze: A discourse perspective. *Language Testing*, 23, 58-72.

- Dick, W., Carey, J. O., & L. Carey. (2000). *The systemic design of instruction*. Upper Saddle River, NJ: Pearson Educational.
- Dupuis, M. M. (1980). *The cloze procedure as a predictor of comprehension in literature*. *Journal of Education Research*, 74(1), 27-33.
- Flech, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- Greene, B.B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading*, 24(1), 82-98.
- Hadley, A. M. (2000). *Teaching language in context*. (3rd ed.). Boston: Heinle & Heinle.
- Halliday, M. A. K., & Hassan, R. (1976). *Cohesion in English*. London: Longman.
- Heaton, J. B. (1998). *Writing English language tests*. NY: Longman.
- Hinofotis, F. B., & Snow, B. G. (1978). *An alternative cloze testing procedure: Multiple-choice format*. In J. W. Oller & K. Perkins (Eds.), *Research in language testing*. Massachusetts: Newbury House.
- Huang, L. (1992). A feasibility study of using the revised cloze test as a measure of reading comprehension. Proceedings of the eighth conference on English teaching and learning in the R.O.C. (pp. 305-332). Taipei: Crane.
- Huang, T. S. (1994). A qualitative analysis of the JCEE English tests. Taipei: Crane.
- Hughes, A. (2003). *Testing for language teachers*. (2<sup>nd</sup>). Cambridge: Cambridge University Press.
- Ito, A. (2004). Two types of translation tests: Their reliability and validity. *System*, 32, 395-405.
- Jonz, J. (1990). Another turn in the conversation: What does cloze measure? *TESOL Quarterly*, 24(1), 61-81.
- Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics*, 17, 81-92.
- Kintsch, W., & Van Dijk, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, 85, 363-394.
- Klein-Braley, C. 1997. C-Test in the context of reduced redundancy testing: an appraisal. *Language Testing* 14: 47-84.
- Kolers, P. A. Some modes of representation. In P. Pliner, L. Krames, and T. Alloway (Eds.), *Communication and effect: language and thought*. New York: Academic Press, 1973.
- Lee, S. H. (2008). Beyond reading and proficiency assessment: The rational cloze procedure as stimulus for integrated reading, writing, and vocabulary instruction and teacher-student interaction in ESL. *System*, 36, 642-660.

- Markman, P. L. (1985). Rational deletion cloze and global comprehension in German. *Language Learning*, 35, 423-430.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276-282.
- Messick, S. (1996). *Validity and washback in language testing*. *Language Testing*, 13, 241-256.
- Miller, M. J., DeWitt, J. E., McCleary, E. M., & O'Keefe, K. J. (2009). Application of the cloze procedure to evaluate comprehension and demonstrate rewriting of pharmacy educational materials. *Annals of Pharmacotherapy*, 43, 650-657
- Oller, J. W. (1979). *Language tests at school*. London: Longman.
- Nicholls, A & Nicholls, S. (1972). *Developing a curriculum: a practical guide*. London: Allen and Unwin.
- Panackal, A. A. & Heft, C. S. (1978). Cloze technique and multiple choice technique: Reliability and validity. *Educational and Psychological Measurement*. 38, 917-932.
- Richards, J. C. & Rodgers, T. S. (2003). *Approaches and methods in language teaching*. (2nd ed.). Cambridge: Cambridge University Press.
- Sasaki, M. (2000). Effects of cultural schemata on students' test-taking processes for cloze tests: A multiple data source approach. *Language Testing*, 17(1), 85-114.
- Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly*, 17(2), 229-255.
- Schmitt, M. C., & Sha, S. (2009). The developmental nature of meta-cognition and the relationship between knowledge and control over time. *Journal of Research in Reading*, 32, 254-271
- Sharp, A. (2009). Reading comprehension in two cultures. *International Journal of Learning*, 16, 281-292.
- Spolsky, B. (1996). *Measured words: The development of objective language testing*. (2nd ed.). Oxford: Oxford University Press.
- Storey, P. (1997). Examining the test-taking process: A cognitive perspective on the discourse cloze test. *Language Testing*, 14, 214-231.
- Stuart, W., & Eve, K. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, 43, 55-77.
- Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.
- Taylor, W. L. (1956). Present developments in the use of the cloze procedure. *Journalism Quarterly*, 33, 42-48.

- Tyler, R. W. (2010). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.
- Viera, A. J. & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine*, 37(5), 360-363.
- Wittrock, M. C. Reading as a generative process. *Journal of Educational Psychology*, 1975, 67, 484-489.
- Wu, H. L. (2002). Rational cloze: Item-generation approaches and construct validity. *English Teaching and Learning*, 26(4), 85-106.
- Yang, T. H. (1996). Fundamental considerations in the test, with special reference to its use in EFL testing in Taiwan. *Sun Yat-sen Journal of Humanities*, 4, 57-77.



## Appendix A

### 教育會考與國中基測比較表

項目	教育會考	國中基測
法源依據	國民小學及國民中學學生成績評量準則	高級中等學校多元入學招生辦法
功能	<p>可使每一位國三學生、教師、學校、家長、主管機關瞭解學生學習品質，並為下一學習階段（高中、高職或五專）作好必要的準備。</p> <p>國中可參酌國中教育會考評量結果，提供學生升學選擇之建議，輔導學生適性入學。</p> <p>可作為高中、高職及五專新生學習輔導參據。</p>	<p>高中、高職及五專多元入學管道的重要依據，主要為測量國三學生各學習領域的基本能力。</p> <p>測驗分數可作為申請入學、甄選入學或登記分發入學的依據。</p>
施測對象	全體國三學生	全體國三學生
辦理時間	預定自民國 103 年起每年 5 月舉辦 1 次	民國 90 年-100 年每年舉辦 2 次； 民國 101 年改為舉辦 1 次
命題依據	國民中小學九年一貫課程綱要能力指標	國民中小學九年一貫課程綱要能力指標
科目	國文、英語（加考聽力）、數學（加考非選擇題型）、社會、自然及寫作測驗	國文、英語、數學、社會、自然及寫作測驗
題型	選擇題與非選擇題型	選擇題型
測驗難度	難易適中	中等偏易
計分方式	標準參照	常模參照
結果呈現	國文、英語、數學、社會及自然評量結果分為精熟、基礎及待加強 3 個等級；寫作測驗分為一至六級分。	除寫作測驗為標準參照之六級分制外，其餘均以量尺分數計算（5 科最高分為 80 分，寫作測驗 12 分，總分為 412）。

資料來源：國立臺灣師範大學心理與教育測驗研究發展中心

國民中學學生基本學力測驗推動工作委員會（<https://cap.nace.edu.tw>）

國民中小學九年一貫課程綱要語文學習領域（英語）

（一）基本理念

隨著地球村時代的來臨，國際間政治、經濟、文化往來頻繁，英語的重要性日益突顯。從資訊、科技、工商業乃至高等教育，英語已成為國際交流的重要溝通工具。此外，透過英語文學習，學習者能適切回應英語國家的社會文化活動和增進多元文化的瞭解與尊重，並進而能使用英語能力為 21 世紀的生活做準備，更被引述為世界公民的必要素質。

為配合政府國際化的政策，提高我國國際競爭力，政府正大力推動各個階段的英語教育，加上社會各界的殷切期盼，英語教學遂規劃九十四學年度起提前至國小三年級開始實施。

國民中小學英語課程旨在奠定國人英語溝通能力的基礎、提升英語學習的動機與興趣、涵詠國際觀、獲致新知，以期未來能增進國人對國際事務的處理能力，增強國家競爭力。本課程強調營造自然、愉快的語言學習環境，以培養學生的學習興趣和基本溝通能力。上課宜採輕鬆活潑、互動的教學模式；教材內容及活動設計宜生活化、實用化及趣味化；體裁宜多樣化。除強調個別語言成分之外，更應強調聽、說、讀、寫四種技能的培養，透過多元教材與活動練習，循序漸進，讓學生藉由同儕及師生的互動，多方面接觸英語，並實際運用，而非由老師單向灌輸文法知識。為了維持學生的學習動機且不增加學習負擔，教材的分量及難易度宜適中，學生的學習興趣與吸收能力應勝於教學進度的考量。

（二）課程目標

國民中小學英語課程的目標為：

- 1.培養學生基本的英語溝通能力，俾能運用於實際情境中。
- 2.培養學生學習英語的興趣與方法，俾能自發有效地學習。
- 3.增進學生對本國與外國文化習俗的認識，俾能加以比較，並尊重文化差異。

（三）分段能力指標

英語科課程分為國小及國中二階段：

第一階段：國小三至六年級。

第二階段：國中一至三年級。

以上二階段的教學同時注重聽、說、讀、寫，不過本課程希望能運用兒童在發音學習方面的優勢，於國小三、四年級的啟蒙階段強調聽、說的學習，讓兒童藉由豐富的英語聽、說的學習，奠定良好的英語口語溝通基礎。然而即使是在啟蒙階段，讀、寫活動並未偏廢，而是適時融入課程，讓學生藉由接觸簡易的閱讀材料，以及適當的臨摹及填寫字詞等練習，自然體驗語言的不同形式，以收聽、說、讀、寫四者相輔相成的效果。第二階段除了持續地培養基本聽說讀寫的能力外，更強調英語的實際運用，發揮其工具性功能，藉以吸收新知、幫助個人身心發展。

本課程的分段能力指標分國小、國中二階段加以規劃，並依語言能力、學習



英語的興趣與方法、文化與習俗三方面加以敘述。其中語言能力的陳述涵括聽、說、讀、寫以及語言綜合應用能力五項。

● 語言能力

讀

階段	分段能力指標
第一階段	3-1-1 能辨識印刷體大小寫字母。 3-1-2 能辨識課堂中習得的詞彙。 3-1-3 能看懂簡易的英文標示。 3-1-4 能辨識歌謠、韻文、故事中的常用字詞。 3-1-5 能看懂簡單的句子。 3-1-6 能辨識英文書寫的基本格式。 3-1-7 能朗讀課本中的對話和故事。 *3-1-8 能藉圖畫、圖示等視覺輔助，閱讀並瞭解簡易故事及兒童短劇中的大致內容。 *3-1-9 能藉圖畫、標題、書名，猜測或推論主題。
第二階段	3-2-1 能辨識英文字母的連續書寫體(cursive writing)。 3-2-2 能用字典查閱字詞的發音及意義。 3-2-3 能看懂常用的英文標示和圖表。 3-2-4 能用適切的語調、節奏朗讀短文、簡易故事等。 3-2-5 能瞭解課文的主旨大意。 3-2-6 能瞭解對話、短文、書信、故事及短劇等的重要內容與情節。 3-2-7 能從圖畫、圖示或上下文，猜測字義或推論文意。 *3-2-8 能辨識故事的要素，如背景、人物、事件和結局。 *3-2-9 能閱讀不同體裁、不同主題的簡易文章。

聽說讀寫綜合應用能力

階段	分段能力指標
第一階段	5-1-1 能正確地辨識、唸出與寫出 26 個英文字母。 5-1-2 能聽懂及辨識課堂中所習得的英語詞彙。 5-1-3 在聽讀時，能辨識書本中相對應的書寫文字。 5-1-4 口語部分至少會應用 300 個字詞，書寫部分至少會拼寫其中 180 個字詞，以應用於簡易的日常溝通中。 5-1-5 能聽懂日常生活應對中常用語句，並能作適當的回應。 5-1-6 能運用字母拼讀法(phonics)。 *5-1-7 能依文字或口語提示寫出重要字詞。

第二階段	5-2-1	能熟習課綱中所標示之 1200 個基本單字，並能應用於日常生活的溝通中。
	5-2-2	能轉述別人簡短的談話。
	5-2-3	能聽懂日常生活對話、簡易故事或廣播，並能以簡單的字詞、句子記下要點。
	5-2-4	能看懂故事及簡易短文，並能以簡短的句子說出或寫出其內容大意。
	*5-2-5	能看懂日常溝通中簡易的書信、留言、賀卡、邀請卡等，並能以口語或書面作簡短的回應。
	*5-2-6	能看懂並能填寫簡單的表格及資料等。





## Appendix C

十二年國民基本教育課程綱要國民中小學暨普通型高級中等學校語文領域-

### 英語文(草案)

#### 壹、基本理念

英語文隨著英語語系國家政治、經濟的龐大影響力，逐漸成為國際通行的共同語言。邁入 21 世紀後，隨著使用人口日漸廣佈全世界，英語似乎也將這個語言的所有權釋放給使用這個語言的國家或個人，容許一定的歧異性，而產生所謂的「世界性英語」(Global Englishes)，這股浪潮進一步讓英語在各地生根茁壯，雖然面對不同語言的挑戰，英語的地位不但屹立不搖，在跨國、跨文化的溝通中，其重要性、普遍性更形日益增強。

在我國，英語並非官方語言，英語文教育是一種外語教育。一般而言，外語課程除了語言知識的學習與語言技能的培養，更應重視語言的使用。我國各級學校的英語文教育也都強調英語文主要是作為人際溝通之用，此理念符合 20 世紀主流的英語文教學理論或教學法。但在充斥巨量資訊且這些資訊多數以英語文的型態流通之際，外語教育有了新的使命與功能，在人際溝通之外，也應培養學生如何藉由外語獲取各個領域的新知，增廣見聞。除了學習語言知識與技能以及他國文化，課程內容也應涵蓋學習方法與策略，以發展學生的自學能力，為未來的終身學習奠定基礎。此外，外語課程也應同時培養思考能力，以期學生在面對知識爆炸的多變環境中，能將資訊化繁為簡，去蕪存菁；進行深層思考，培養

多元觀點，一方面尊重多樣性的文化，並珍惜自然環境與資源，以期永續發展；一方面豐富個人視野，增加創意、彈性與適應力。在發展上述諸多認知能力之餘，切勿忘記學習興趣的培養、學習動機的激發，而且，學生的學習動機及實際的消化吸收應優於教學進度的考量。尊重學生的個別差異與需求，提供適性學習的環境，落實學習機會上的公平正義。簡言之，我國的英語文課程與教學應該具有以下特質：

- 一、以學生為中心，重視適性學習的原則與學習者的情意因素。
- 二、重視語言溝通互動的功能性，凸顯其獲取新知的工具性角色。
- 三、發展學生自主學習與終身學習英語文的能力與習慣。
- 四、引導學生獨立思考、處理與運用訊息的能力。
- 五、透過語言學習探索異國文化，進行跨文化反思，提升社會參與並培養國際觀。
- 六、培養學生邏輯思考能力，並激發其創造力。

## 貳、課程目標

根據以上基本理念，英語文課程涵蓋以下幾個目標：

- 一、培養英語文聽、說、讀、寫的能力，應用於日常生活溝通。
- 二、提升學習英語文的興趣並涵育積極的學習態度，主動涉獵各領域知識。
- 三、建構有效的英語文學習方法，強化自學能力，奠定終身學習之基礎。
- 四、尊重與悅納多元文化，培養國際視野與全球永續發展的世界觀。

五、培養以英語文進行邏輯思考、分析、整合與創新的能力。

● 語言能力（讀）

階段	學習表現
第二學習 階段	3-II-1 能辨識 26 個印刷體大小寫字母。 ◎3-II-2 能辨識課堂中所學的字詞。 ◎3-II-3 能看懂課堂中所學的句子。
第三學習 階段	◎3-III-1 能辨識課堂中所學的字詞。 3-III-2 能看懂校園內簡易的英文標示。 3-III-3 能看懂教室用語。 3-III-4 能看懂課堂中所學的句子。 3-III-5 能看懂課堂中所學的簡易對話。 3-III-6 能看懂課堂中所學的簡易短文之主要內容。 * 3-III-7 能看懂繪本故事的主要內容。 * 3-III-8 能看懂兒童短劇中的主要內容。 * 3-III-9 能藉圖畫、標題、書名等作簡易的猜測。
第四學習 階段	3-IV-1 能辨識連續書寫體大小寫字母。 ◎3-IV-2 能辨識課堂中所學的字。 3-IV-3 能看懂簡易的英文標示。 3-IV-4 能看懂簡易的圖表。 3-IV-5 能看懂簡易的生活用語。 3-IV-6 能看懂基本的句型。 3-IV-7 能了解對話的主要內容。 3-IV-8 能了解短文、簡訊、書信的主要內容。 3-IV-9 能了解故事的主要內容與情節。 3-IV-10 能辨識簡易故事的要素，如背景、人物、事件和結局。 3-IV-11 能藉圖畫、標題、書名等作合理的猜測。 * 3-IV-12 能熟悉重要的閱讀技巧，如擷取大意、猜測字義、推敲文意、預測後續文意等。 * 3-IV-13 能了解短劇的主要內容與情節。 * 3-IV-14 能快速閱讀了解文章重點，並有效應用於廣泛閱讀中。 *◎3-IV-15 能分析及判斷文章內容，了解敘述者的觀點、態度及寫作目的。 * 3-IV-16 能閱讀不同體裁、不同主題的簡易文章。
第五學習 階段	◎3-V-1 能辨識課堂中所學的字詞。 3-V-2 能看懂常見的英文標示。 3-V-3 能看懂常見的圖表。 3-V-4 能看懂常見的生活用語。 3-V-5 能看懂常見的句型。 3-V-6 能了解對話的內容。

	3-V-7 能了解短文、書信的內容及文本結構。 3-V-8 能了解故事的內容與情節。 3-V-9 能了解短劇的內容與情節。 3-V-10 能辨識故事的要素，如背景、人物、事件和結局。 3-V-11 能閱讀不同體裁、不同主題的文章。 3-V-12 能利用字詞結構、上下文意、句型結構及篇章組織推測字詞意義或句子內容。 3-V-13 能熟悉各種閱讀技巧（如擷取大意、推敲文意、預測後續文意），進行快速閱讀並有效應用於廣泛閱讀中。 3-V-14 能從圖畫、圖示或上下文，猜測字義或推論文意。 ◎3-V-15 能分析及判斷文章內容，了解敘述者的觀點、態度及寫作目的的。
--	--

● 語言能力（聽說讀寫綜合應用能力）

階段	學習表現
第二學習階段	5-II-1 能正確地認讀與聽寫 26 個字母。 ◎5-II-2 在聽讀時，能辨識書本中相對應的書寫文字。 ◎5-II-3 能以正確的發音及適切的速度朗讀簡易句型的句子。 ◎5-II-4 能運用所學的字母拼讀規則讀出英文字詞。
第三學習階段	5-III-1 能認讀與聽寫國小階段字詞。 ◎5-III-2 在聽讀時，能辨識書本中相對應的書寫文字。 5-III-3 能聽懂、讀懂國小階段基本字詞及句型，並使用於簡易日常溝通。 ◎5-III-4 能聽懂日常生活應對中常用語句，並能作適當的回應。 ◎5-III-5 能以正確的發音及適切的速度朗讀簡易句型的句子。 5-III-6 能以正確的發音及適切的速度朗讀或吟唱簡易歌謠韻文。 5-III-7 能以正確的發音及適切的速度朗讀簡易對話。 5-III-8 能以正確的發音及適切的速度朗讀簡易故事及短劇。 ◎5-III-9 能運用所學的字母拼讀規則讀出英文字詞。 * 5-III-10 能運用所學的字母拼讀規則拼寫英文字詞。
第四學習階段	5-IV-1 能聽懂、讀懂國中階段基本字詞，並使用於簡易日常溝通。 5-IV-2 能掌握國中階段所學字詞及句型，適當地使用於日常生活之溝通。 ◎5-IV-3 能聽懂日常生活應對中常用語句，並能作適當的回應。 5-IV-4 能以正確的發音及適切的語調及速度朗讀短文及短劇。 * 5-IV-5 能運用字母拼讀規則讀出及拼寫英文字詞。 5-IV-6 能轉述所聽到的簡短談話。 5-IV-7 能聽懂日常生活對話，並能以簡單的字詞、句子記下要點。 5-IV-8 能聽懂簡易故事，並能以簡單的字詞、句子記下要點。 5-IV-9 能聽懂簡易廣播，並能以簡單的字詞、句子記下要點。

	<p>5-IV-10 能讀懂簡易故事及短文，並能以簡短的句子說出或寫出其內容大意。</p> <p>5-IV-11 能看懂並能填寫簡單的表格及資料等。</p> <p>* 5-IV-12 能看懂日常溝通中簡易的書信、簡訊、留言、賀卡、邀請卡等，並能以口語或書面作簡短的回應。</p>
第五學習 階段	<p>5-V-1 能聽懂、讀懂高中階段基本字詞，並使用於日常溝通。</p> <p>5-V-2 能掌握高中階段所學字詞及句型，適當地使用於日常生活之溝通。</p> <p>5-V-3 能以正確的發音及適切的斷句、節奏、語調、語氣及速度，流暢地朗讀短文、短劇及故事。</p> <p>5-V-4 能針對各類選文，以口語或書面回答相關問題。</p> <p>5-V-5 能以自己的話轉述一段談話或簡短故事。</p> <p>5-V-6 能看懂並填寫常用的表格。</p> <p>5-V-7 能聽懂日常對話，並能記下要點。</p> <p>5-V-8 能聽懂故事，並能記下要點。</p> <p>5-V-9 能聽懂廣播，並能記下要點。</p> <p>5-V-10 能讀懂故事及短文，並以簡短的句子述說或寫出主旨或大意。</p> <p>5-V-11 能看懂日常書信、簡訊、電子郵件、留言和賀卡、邀請卡等，並能以口語或書面作回應。</p> <p>* 5-V-12 能以適切的英語文說出或寫出談話或短文的摘要。*</p> <p>5-V-13 能了解及欣賞不同體裁、不同主題之文章，並據以發表心得或感想。</p>

## Appendix D

The Distribution of the Passing Rates of the Cloze items in CAP from 2014 to 2017

Years	Question Numbers	Passing Rates
2014	33	0.4
	34	0.44
	35	0.51
	57	0.33
	58	0.55
	59	0.41
	60	0.46
2015	13	0.57
	14	0.53
	15	0.63
	16	0.4
	17	0.55
	18	0.5
	35	0.52
2016	36	0.48
	37	0.54
	38	0.32
	39	0.55
	40	0.47
	41	0.38
	34	0.5
2017	35	0.49
	36	0.63
	37	0.59
	38	0.42
	39	0.55
	40	0.45
	41	0.51

Source: Research Center for Psychological and Educational Testing at National Taiwan Normal University. (<https://cap.nace.edu.tw>)



## Appendix E

### The cloze test section of 2010-1 BCT

#### 二. 題組：

(19-21)

Mr. Hutman owns a restaurant. One day, one of his two cooks ran away with some money. Mr. Hutman was sad and worried because the next day was Saturday, and the restaurant 19. very busy then. So that night he called his friends, but could not find anyone to help.

On Saturday, many people came to the restaurant for lunch. 20. cook got so busy that he wanted to leave, too. 21., one of the waiters, Henry, told Mr. Hutman that he could help in the kitchen. People who had lunch in the restaurant that day enjoyed the food Henry cooked. Mr. Hutman was very happy and made Henry a new cook of the restaurant.

- (A) 19. (A) would be (B) is (C) has been (D) will be  
(D) 20. (A) One (B) Each (C) Another (D) The only  
(C) 21. (A) Also (B) In fact (C) Luckily (D) For example

(22-24)

(In an English class)

Teacher : Let's play a game with 3-letter words. The game is easy to play. First, a 3-letter word is spoken. Then the next person has to say a different 3-letter word that starts with the last letter of the spoken word. Anyone who cannot speak a right 3-letter word in two seconds will have to 22. Any questions?

Brian : Yes. Can we try again if we say a wrong word?

Teacher : No, you 23. That's why the game is exciting and fun. Now let's try it!

Brian : OK! Let me try first. MAP.

Amy : PUT.

Jane : TEA.

Ken : 24.

Carl : EAR.

Sam : RAIN.

Carl : Ha! I got you.

Sam : Oh, no! I forgot it has to be a 3-letter word.

Teacher : Sam, it's your show time. Here's the book. Pick out any story you'd like. Now everybody, let's listen to Sam.

- (B) 22. (A) sing a song (B) read a story  
(C) draw a picture (D) make a funny face  
(A) 23. (A) have just one chance  
(B) have to give a different word  
(C) may look it up in the dictionary  
(D) must ask your classmates for help  
(D) 24. (A) TEN (B) RED (C) EAT (D) AGE

二. 題組：

(17-20)

People in Quiet Town are very sad that their old friend, Time River, is facing the end of its life next October. The river 17. a lot to the people in the town. They all grow up and live with it all their lives.

In the old times, people in the town 18. their houses on both sides of the river. The houses in each block shared a public washing room. On Mondays, people came and washed their clothes together there. They hung their clothes at the back of each house. Those clothes with different colors 19. a beautiful picture along the river.

Now Time River is going to be filled in to make more land to build on. People are happy that they 20. more living space. But at the same time, they feel sorry that they must say good-bye to their old friend. So they have decided to give a party one day a year to remember the river.

📖 hung (原形為 hang) 懸掛

- |                  |                  |                |           |
|------------------|------------------|----------------|-----------|
| (A) 17.(A) means | (B) is meaning   | (C) will mean  | (D) meant |
| (D) 18.(A) build | (B) are building | (C) have built | (D) built |
| (D) 19.(A) make  | (B) are making   | (C) have made  | (D) made  |
| (C) 20.(A) have  | (B) have had     | (C) will have  | (D) had   |

(21-23)

A small town has a good chance of 21. that can bring in a lot of money, if it has something special to be proud of. One example is Gukeng town of Yunlin, Taiwan. Gukeng has long been famous for growing good coffee, but the town didn't start to make much money from it until some years ago. As more and more people have visited Gukeng for its coffee, the coffee farmers have begun to open their farms to the public. At these farms, people can have the fun of finding out where coffee comes from. 22., coffee shops are opened all over Gukeng, and people can take a rest and taste delicious coffee on the sidewalks in or after a day's visit. The new businesses make a better life possible for those who 23. the town. They don't have to leave the town to find jobs in other places.

📖 as 隨著

- |                                 |                               |
|---------------------------------|-------------------------------|
| (B) 21.(A) growing the best tea | (B) starting a new business   |
| (C) selling old farming lands   | (D) opening a shopping center |
| (B) 22.(A) First                | (B) Also                      |
| (C) However                     | (D) For example               |
| (A) 23.(A) live in              | (B) hear about                |
| (C) take a trip to              | (D) are interested in         |



The cloze test section of 2011-1 BCT

二. 題組：

(21-24)

Long long time ago, a girl named Doris lost her ring. She felt so bad that she could not sleep well. When Doris turned over in her bed, she found the black horse in the picture on the wall \_\_\_\_21\_\_\_\_. She was very surprised, so she got out of bed to have a closer look. When Doris touched the head of the horse, she \_\_\_\_22\_\_\_\_ into the picture!

Now Doris was sitting on the back of the running horse! It was so exciting to be riding a horse in the beautiful sky, but Doris did not know where the horse \_\_\_\_23\_\_\_\_ her. Finally, they stopped in front of an old church. There Doris found a gift bag. She opened the bag and saw her ring inside. Doris was very happy. When she put on the ring, she woke up from her dream. At that moment, Doris could not believe her eyes – the ring \_\_\_\_24\_\_\_\_ there back on her finger!

- (C) 21. (A) is moving (B) has moved (C) was moving (D) will move  
(B) 22. (A) is pulled (B) was pulled (C) has been pulled (D) would be pulled  
(D) 23. (A) took (B) is taking (C) has taken (D) would take  
(B) 24. (A) is (B) was (C) has been (D) would be

(25-26)

Mr. Lee : Wow! You look great in \_\_\_\_25\_\_\_\_ glasses! Its glasses are very famous. How much is this pair?

Mr. Kao : Only 500 dollars.

Mr. Lee : Really? Where did you get the glasses?

Mr. Kao : At B. B. Department Store. It's having a sale these days.

Mr. Lee : Does the department store have any \_\_\_\_26\_\_\_\_ on sale, too? I want to buy a pair for my wife's birthday.

Mr. Kao : Yes, it does. I remember I have the flyer. Here. You'll get a much better price if you buy three pairs at a time.

Mr. Lee : Thanks.

B. B. DEPARTMENT STORE  
Special Prices This Week

**Glass House**  
glass plate / dish / bowl  
any three pieces for \$1,500

**Rian's**  
■ glasses 1 for \$900  
■ blue jeans  
1 for \$800  
3 for \$1,800

**O.L.S. Glasses**  
two pairs for \$500

**NUMA**  
\* sports shoes  
- one pair for \$700  
- two pairs for \$1,200  
\* glasses \$500/pair  
\* tennis shorts \$300/pair

📖 flyer 廣告傳單

- (D) 25. (A) Glass House (B) O.L.S (C) Rian's (D) NUMA  
(B) 26. (A) glass plates (B) blue jeans (C) sports shoes (D) tennis shorts

二. 題組：

(21-24)

My mother likes to tell me the story about my name, Stella. She 21. about it many times. I was born on a night full of stars. My mom 22. a hard time when the doctor said I was too big to be pulled out. At that moment, Mom looked into the sky, making a wish with all her heart, "Let the baby live, please, even if I may die!"

Just then, a very bright star 23. from the sky. All her troubles were soon gone, and my little head came out of her body. With the falling of a star came a new life. That's how I 24. the name "Stella," which means "star."

(C) 21. (A) talked (B) is talking (C) has talked (D) will talk

(A) 22. (A) was having (B) has (C) has had (D) will have

(B) 23. (A) would fall (B) fell (C) falls (D) is falling

(A) 24. (A) got (B) get (C) am getting (D) will get

(25-27)

Could it be possible to live in a house made of ice? What is it like to sleep on an ice bed? **North Hotel** is the magic place that brings people from all over the world to 25. in the cold. Here you can have dinner at an ice table, watch movies in an ice chair, and even take a bath in a large ice box. Worried about catching a cold? 26. ! **North Hotel** keeps you warm and makes you feel comfortable all the time. You can live in this dreamland for a few days and enjoy many special experiences in the snow. It is not easy to tell you about everything here at **North Hotel**; you must 27. . Pack now and get ready to be COOL!

(A) 25. (A) have fun (B) start a new life

(C) get some exercise (D) learn to make friends

(B) 26. (A) Good idea (B) You don't have to be

(C) Take care of yourself (D) Don't forget your heavy coat

(B) 27. (A) think about it carefully

(B) come to see it for yourself

(C) check out the weather first

(D) wait and see what will happen

二. 題組：

(21-23)

Dear Jean,

How have you been? I'm writing this to say I'm sorry that 21. your birthday. These days I felt there was something wrong, but I just couldn't think of what it was. Yesterday afternoon on the bus, I 22. remembered what it was. Your birthday is this month, and it was two weeks ago! I feel so bad. I picked out a gift for you. I hope you like it.

By the way, 23. a great time on your birthday? I can't wait to hear about it.

Love,

Linda

📖 by the way 順便一提

(B) 21.(A) I forget (B) I forgot (C) I'm forgetting (D) I've forgotten

(B) 22.(A) also (B) finally (C) only (D) still

(D) 23 (A) have you had

(B) are you having

(C) would you have

(D) did you have

(24-26)

If you sail between Hawaii and California, you'll find a big part of the ocean, as big as Africa, covered with garbage. Most of the trash comes from the land, by wind or by water, and some is thrown straight into the sea. Most of the trash is plastic : plastic bags, plastic cups, plastic cans, and a lot more.

24. ? Some just stays in the sea, some is washed onto the beaches, and, worst of all, some 25. . Many seabirds and sea animals that eat the trash in the sea die. Those that are lucky enough to live on may someday arrive on our dinner plate. What we throw away will all 26. one way or another.

Keep in mind what happens to trash, and you may think twice before you throw anything away.

📖 ocean 海洋

(A) 24.(A) Where does the trash go

(B) How do we deal with the trash

(C) Where does the trash come from

(D) How does the trash arrive in the sea

(A) 25.(A) enters seabirds' and sea animals' stomachs

(B) stops seabirds and sea animals from growing up

(C) fills the living space of seabirds and sea animals

(D) goes down to the bottom of the sea and never comes up

(B) 26.(A) be cleared up (B) come back to us

(C) be useful to others (D) get out of our sight

二. 題組：

(19-21)

Dear Sir,

I'm sorry that I have to write this letter, but if I 19., you'll never see what your problem is.

On Saturday evening, my family and I went to your restaurant for my birthday. Many people 20. in the restaurant. We were very hungry and didn't want to wait, but your waiter said we 21. served in twenty minutes. We believed him, but we ended up waiting for ONE HOUR!

I have to say the steak was delicious, but it was still wrong to keep people waiting for so long.

Julia Wu

📖 serve 服務

(B) 19.(A) didn't (B) don't (C) hadn't (D) haven't

(A) 20.(A) were already (B) are already  
(C) have already been (D) will already be

(B) 21.(A) were (B) would be (C) have been (D) are going to be

(22-24)

Gisele : My dad's taking me to play the Dye-Out game this Saturday. He asked if you would like to come.

Opal : 22.

Gisele : It's a popular game these days. The game is played by two teams and can be played inside or outside. You win the game by taking the other team's castle or "killing" all the players on the other team.

Opal : 23. ?

Gisele : You use balls filled with oil paint to attack the players on the other team. Their clothes will be covered with paint when they are hit by the balls. This way, they are "killed" and have to leave the playground.

Opal : Sounds exciting.

Gisele : It is, but 24. For example, you have to buy the clothes, the balls, the masks, and many other things. Also, it can be dangerous sometimes.

Opal : Well, then I have to ask my parents first.

(C) 22.(A) But I'm not good at it

(B) Your dad has already invited me

(C) But I've never heard of it before

(D) I'd love to join you this Saturday

(B) 23.(A) How do you take the castle

(B) What do you mean by "killing" a player

(C) What do you do after you "kill" a player

(D) How do you know where the players are hiding

(A) 24.(A) it's an expensive hobby

(B) it hurts when you get hit

(C) you need a big space for the game

(D) your clothes will get dirty with the paint

二. 題組：

(33-34)

Do you eat meat? Well, if you do, then you might find our next news interesting. A study says that the world has a 33. need for meat. In 1960, the world ate 64 million tonnes of meat, about 21 kg for each person. In 2007, the number rose to 268 million tonnes, about 40 kg for each person. At the same time, 34. In the 1960s, beef was high on the menu. Of the meat that was eaten, 40% was beef. In 2007, pork became the star. Poultry also became popular, going up from 12% to 32%, thanks to people's worries about their health these years. And 35.? You're guessing the U.S.A., right? The answer is Luxembourg! In 2007, each Luxembourger ate about 137 kg of meat! Second to Luxembourgers are Americans. In 2007, each American ate about 126 kg! Now, enough with the numbers. I'm playing you a song called *Currywurst*. The singer sings about his love for the meat dish of the same name. Enjoy!

tonne 公噸 poultry 家禽 (肉)

- (C) 33.(A) falling (B) special (C) growing (D) common  
(C) 34.(A) we have changed our way of cooking meat  
(B) new kinds of meat have come on the market  
(C) there have been changes in the list of favorite meats  
(D) doctors have been worried that we eat too much meat  
(D) 35.(A) where does the world's best meat come from  
(B) which country is the world's biggest meat maker  
(C) where can you eat the world's most delicious meat  
(D) which country uses up the most meat for each person

(57-60)

**Reader's Story**

It's never comfortable to take a bus during rush hour on a hot summer day. All my friends hate it. And I..., well, I 57. it, too.

The story happened two months ago. I was on a bus that was packed with almost everyone in the city. I 58. my friend Neal to see a show. And then he called and said he couldn't

come. "This is not what I planned for Friday!" I thought. The air on the bus was terrible; the man in front of me smelled like dead fish. While I was feeling sorry for myself, a girl called my name. I didn't recognize her at first. Then I was surprised to find that she was Hui-ting, my old neighbor. I 59. Hui-ting in years. We used to play together every day when we were kids. We were happy to see each other, so we decided to have dinner together. And that was the start of many dinners after.

With Hui-ting, the bus ride 60. something I love. The bus is still packed with people, but I enjoy the ride every day.

(Xiang Chang, Taipei)

rush hour 尖峰時間 recognize 認出

- (C) 57.(A) hate (B) have hated (C) used to hate (D) would hate  
(B) 58.(A) used to meet (B) was going to meet (C) have met (D) had met  
(D) 59.(A) won't see (B) wouldn't see (C) haven't seen (D) hadn't seen

( A ) 60.(A) has become (B) had become (C) will become (D) would become



二. 題組：

(13-15)

From : Lisa Clyne (lisaclyne@mail.com)

To : Mary Faber (maryfaber@mail.com)

Date : Thursday, April 2, 2015

Subject : Welcome

-----  
Dear Mary,

Uncle Billy and I are excited about your visit.

The weather has \_\_13\_\_ these days. Like today, it was sunny in the morning, but rained heavily at noon. So we've \_\_14\_\_ : If the weather is nice, I'm going to take you to Smith Farm. They're having a special horse show this week.

I'm sure you'll like it. And \_\_15\_\_. We can go shopping at Mimi's

Department Store; there we can walk around without getting wet. Rain or shine, we hope you will have a good time here.

See you tomorrow at the train station at 9 : 00 a.m.

Love,

Aunt Lisa

( C ) 13.(A) got better

(B) been warm

(C) changed a lot

(D) become wet and cold

( A ) 14.(A) made two plans for you

(B) worried about your trip to here

(C) decided to take you to a nice place

(D) prepared everything you've asked for

( B ) 15.(A) we'll agree with each other

(B) don't worry if it rains

(C) we all love animals very much

(D) don't forget to check the weather

(16-18)

**NEWS HUNT**

Kieran Hardy 11/07/2013

For years, we thought our Earth was the only blue dot up there. Now another has been found. Its name is HD189773b. HD189773b is an exoplanet, a planet outside our Solar System, and is one of the nearest exoplanets to Earth. Even so, \_\_16\_\_. Here's why : It is 63 light years away.

That means it is 370,440,000,000,000 miles from us. Even if we fly at 3,500 miles an hour, it will take more than 12 million years to get there.

\_\_17\_\_ : It is much bigger, it is made of gas and it is burning hot. In heat as great as 1,000°C, life is not possible. What is worse, it rains glass. If 1,000°C does not kill you, glass rain will.

Even so, finding this blue giant \_\_18\_\_. It is the first time that we have been able to see the color of an exoplanet. The color of a planet gives us ideas about what is happening on it. While we have a long way to go before

we find a new planet to live on, finding a blue dot is a good start.

📖 exoplanet 系外行星 Solar System 太陽系

- (A) 16.(A) we cannot really call it a neighbor  
(B) we do not know anything about it yet  
(C) we might not be able to stay there for long  
(D) we are not sure how long it takes to get there
- (C) 17.(A) What's more, it is not water that makes HD189773b look blue  
(B) With its blue color, HD189773b could be a second Earth for us  
(C) Except for its blue color, HD189773b is nothing like our home planet  
(D) Because of its blue color, people guess there might be life on HD189773b
- (A) 18.(A) gives us hope  
(B) took hard work  
(C) has changed our life  
(D) helps us know more about Earth





二. 題組：

(35-38)

It was 11 p.m. and Molly walked out of her bakery. She turned and looked at her store one last time. She wanted to remember what it 35. at the moment. A few hours later, people would come and clean out everything in the store. A young man had bought it. He 36. it into a flower shop.

Before it was a bakery, this place 37. a small coffee shop. Molly worked in the shop as a waitress. But taking orders was never Molly's dream; baking was. When she knew her boss planned to sell the shop, she borrowed money and bought it.

Her bakery had been open for thirty years. Thirty very wonderful years. However, it would all come to an end tonight. Molly 38. the bakery to be a family business. But her daughter was never interested in baking. Molly did not want her business in a stranger's hands, so after some serious thinking, she decided to close it.

"Goodbye, my dear old friend," Molly looked at the store, whispering.

📖 whisper 低語

- (B) 35.(A) had looked like (B) looked like  
(C) would look like (D) has looked like  
(C) 36.(A) had changed (B) changed  
(C) was going to change (D) has changed  
(A) 37.(A) used to be (B) would be (C) has been (D) is  
(A) 38.(A) had wanted (B) has wanted (C) would want (D) will want

(39-41)

What does the word *family* mean to you? An American study in 2006 showed that people today 39. Over 99% of the people who were interviewed agree that a husband, a wife, and a child are a family. At the same time, 94% see a parent with a child as a family, 91% say a husband and a wife, without children, are a family, and 81% think a man and a woman, with a child, but not married, are a family too.

The study also found that 40. is very important in the modern thinking on family. Though 81% think a man and a woman, not married, with a child, are a family, the percentage (%) drops to 40% if the couple doesn't have a child. This is also true with same-sex couples. About 60% see two men, or two women, with a child, as a family, but only 32% think so when the couple doesn't have a child.

In the study, those who see two men or two women that live together as a family often find it OK for same-sex couples to get married. 41. However, not everyone opens their arms to same-sex couples: the study said 30% have no problem seeing pets as part of one's family, but they do not think a same-sex couple is a family.

📖 couple 伴侶 same-sex 同性

- (D) 39.(A) think differently about when to start a family  
(B) do not find family as important as their parents did  
(C) want many different things when they start a family

- (D) have several different ideas about what makes a family
- ( B ) 40.(A) whether people are married or not  
(B) whether people have a child or not  
(C) whether people live together or not  
(D) whether people love each other or not
- ( A ) 41.(A) This is not surprising  
(B) This is not possible everywhere  
(C) It is no good news for everyone  
(D) It cannot be this way for very long



二. 題組：

(34-37)

Jerry Stevens has been unhappy these months. He 34. a proud businessman of a successful shaved ice shop. His highest sales numbers were 1,899 plates a week; people called him “Ice King.” But all this changed when the cookie shop across the street started selling ice cream cookies. It took away half of his business.

And 35. when Dan, his dead sister’s only son, came to work for him.

Dan was a nice young man with a simple mind, too simple, maybe. He always let people try some shaved ice for free, but he never learned to see when “some” became “too much.” Many people came, but few bought. When the sales did not grow with the heat, Jerry started to worry. He wanted to send Dan home. But how could he? Dan had no family except him.

Then, one day, 36. And it was from across the street. The cookie shop 37. used to look for a clerk. For Jerry, it was a great chance to get Dan out of his store without sending him home. After taking the shopkeeper to several nice dinners, Jerry got Dan the job.

Now Jerry could finally try and save his store. And maybe Dan would help by sending him a few people who are thirsty from eating too many cookies.

( C ) 34. (A) is (B) has been (C) used to be (D) was going to be

( A ) 35. (A) things didn’t get better  
(B) he was ready to give up  
(C) the sales of shaved ice finally started to rise  
(D) the cookie shop was not successful for long

( B ) 36. (A) his only hope died (B) good news arrived  
(C) worse trouble came (D) the hard time passed

( C ) 37. (A) used to look for (B) has looked for  
(C) was looking for (D) had looked for

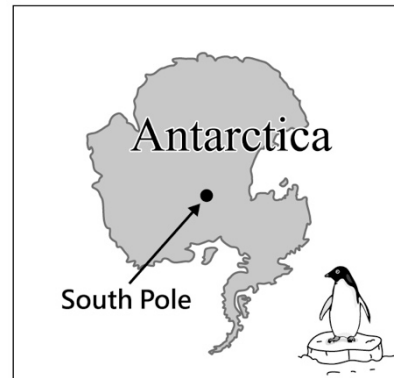
(38-41)

Englishman Robert Scott is known for leading two trips to Antarctica. The first one made him a star; the second  
38. \_\_\_\_\_.

After his first successful trip, Scott decided to be the first person to stand on the South Pole. However, 39. \_\_\_\_\_. He would be in a race with Roald Amundsen, from Norway. Both left their countries by ship in June of 1910 and arrived in Antarctica in January of 1911.

About ten months later, both teams started their trips down to the South Pole.

40. \_\_\_\_\_. Amundsen's team used dogs, and Scott's team used horses. Because horses weren't good at traveling on snow, it took Scott's team 77 days to arrive at the South Pole. They got there on January 17, 1912, and were surprised to find that Amundsen was ahead of them. The news hit Scott very hard, but what he didn't know was that 41. \_\_\_\_\_. His team began their long trip home with broken hearts. After days of terrible weather and little food, Scott lost his men one after another, and he himself was the last one to meet the end of his life. No one on his team lived to go back home and tell their story. It was only learned through Scott's diary.



- (C) 38. (A) hurt his health (B) opened his eyes  
(C) cost him his life (D) made his dream come true
- (B) 30. (A) he didn't want to take this trip with others  
(B) he wasn't the only one who had this dream  
(C) he wouldn't let anyone else take his prize away  
(D) he couldn't miss the chance to make his country proud
- (D) 40. (A) And Scott used a popular way for his men to travel fast  
(B) But there were problems to deal with before they traveled  
(C) But Amundsen didn't know better about animals than Scott  
(D) And their ways of traveling decided which team would win
- (C) 41. (A) there was still hope (B) no one ever believed him  
(C) the worst had not come yet (D) Amundsen had not been honest