Investigating the Washback Effect of the New CET 4 Listening Comprehension Subtest on Language Learners

WANG Weili
Northwest Agriculture & Forestry University

Abstract

The new CET 4 LCS is introduced with the intention of giving impetus to college English teaching and improving non-English major college students’ listening comprehension. The present study set out to investigate the washback effect of the new CET 4 LCS on language learners. Based on Hughes’ washback model (1993) and empirical washback studies, this study focused on learner participants’ attitudes toward aspects of the new CET 4 LCS and their corresponding behavior influenced by the listening subtest in their learning process. Data were collected from 329 non-English major undergraduates of 34 colleges and universities in Xi’an by means of a questionnaire survey. Results indicate that there are both positive and negative washback effects in students’ attitudes and behaviors.

Key words: washback effect; new CET 4 LCS; learners

I. Introduction

The new College English Test Band Four (CET 4) is designed in accordance with the new college English teaching syllabus, i.e., the College English Curriculum Requirements (CECR), with the aim of measuring students’ overall English proficiency, promoting the implementation of the new teaching syllabus and bringing about beneficial washback effects on teaching and learning. Similarly, the new CET 4 Listening Comprehension Subtest (LCS) is aimed at measuring students’ listening comprehension ability and promoting good teaching and learning practice of listening comprehension. A shift from focus on reading and writing to focus on listening and speaking is highly emphasized in this use-oriented approach to testing.
The CET 4 certificate gained social recognition as a measure of the quality of English learning and teaching and had considerable impact on students’ graduation and even their future career (Jin, 2007). Although this certificate has been abandoned and replaced by a score report, performance on this test still has serious consequences and implications for various stake-holders, due to its high-stakes and its “social weight” (Gu, 2007: 1). As a result, the new CET 4 can be expected to exert strong influence on teaching and learning, which is often defined as washback effect in applied linguistics. The focus of the CET 4 test designers has shifted from reliability to validity and from validity to beneficial washback effects in recent years (Gu, 2007: 2). Besides, “CET has had profound impact on the quality of College English teaching” (Gu, 2007: 46). Additionally, the CET 4 LCS has undergone the greatest change in terms of its test formats and score percentage. Consequently, it is imperative to investigate its washback effects.

“Definitions of washback are nearly as numerous as the people who wrote about it” (Bailey, 1999: 3). Researchers have adopted different perspectives as to its definition and scope according to their specific research context and target population. In view of this, the term washback is preferred in this study to indicate the influence of the new CET 4 LCS on college students’ learning in Chinese educational context. In short, the present study attempts to investigate the washback effect of the new CET 4 LCS on language learners, for the purpose of examining how a high-stakes subtest influences students’ learning in Chinese educational context.

2. Literature Review

2.1 Washback Models

There are various washback models. Among them, the traditional model by Alderson and Wall (1993) includes a number of washback hypotheses, and Hughes’s (1993) model distinguishes between participant, process and product.

Alderson and Wall’s washback hypotheses (1993) provide the basic criteria for the research design in this study and Hughes’ basic washback model (1993) is used, in a simplified way, as a framework for analysis in this study. According to this model, “the nature of a test may first affect the perceptions and attitudes of the participants, which in turn may affect what the participants do in carrying out their work (process), which will affect the learning outcomes, the product of that work” (Hughes, 1993: 2). It emphasizes “participants’ perceptions and attitudes and how these factors affect what they do” (Bailey, 1999: 10). These aspects, therefore, are the research foci of this study.

2.2 Learner Washback Studies

Compared with teacher washback studies, learner washback studies are still limited in number, and are always conducted as part of broader studies. The studies by Watanabe (1992, 1996), Shohamy et al. (1996) and Andrew et al. (2002) show that tests do influence learners, and that individual learners will experience this influence in different ways; that is, “the nature of washback seems to vary from student to student” (Andrew et al., 2002:
207). There are also studies that focus on the relationship between teacher perceptions and learner expectations such as Alderson and Hamp-Lyons (1996), Hays and Read (2004), Green (2006), Chapman (2007) and Tsagari (2007).

Learner washback studies give great inspiration to the present study. They show that tests do influence learners and that it is important to look into learner attitudes and behaviors.

3. Research Methodology

The following two research questions were explored in this study to examine the nature of the washback effect of the new CET 4 LCS on language learners.

Research Question One: What is the nature of the washback effect on students’ attitudes toward the new CET 4 LCS?

Research Question Two: What is the nature of the washback effect on students’ learning behaviors as a result of the new CET 4 LCS?

3.1 Participants

Participants in this study included 329 second and third year non-English major undergraduate students from 34 colleges and universities in Xi’an, Shaanxi Province. Among them, 112 were from Xi’an Institute of Finance and Economy, 39 were from Shaanxi University of Science and Technology, and 37 were from Xi’an Polytechnic University. Participants from these three universities made up 64.2% of the total research sample. The remaining 35.8% were from 31 other colleges and universities in Xi’an, with the number of participants in each college or university ranging from one to eleven.

A questionnaire was delivered to all the 329 participants with a 100% return rate, and among them there were 293 valid responses and 36 invalid ones. Among the 293 participants who had finished a valid questionnaire, 200 did not have CET 4 experience and were likely to take it at the end of the following semester, i.e., the fourth semester, due to the requirements of their universities, while the other 93 students had at least one test experience.

3.2 Instruments

The primary instrument of the present study was a questionnaire written in Chinese, and items in the questionnaire were based on the following sources:

1) Alderson & Wall’s fifteen washback hypotheses (1993)
2) Bailey’s washback model (1996)
3) CET documents and past test papers
4) Cheng’s HKCEE washback study (2005)
5) College English Curriculum Requirements (2007)
6) College English Test Band Four syllabus (2007)
7) Gu’s CET washback study (2007)
8) Hughes’ basic washback model (1993)
9) Qi’s NMET washback study (2005)

The design and validation procedure of the questionnaire underwent four main stages: 1) An initial draft based on the nine sources listed above; 2) Iterative revisions based on informal conversations with non-English major college students about their attitudes toward the new CET 4 LCS and learning behaviors influenced by it; 3) Three round pilot testing with second-year non-English major college students at Shaanxi University of Science and Technology in June, 2007, with each round using a revision questionnaire based on the previous piloting results; 4) A further revision of the questionnaire based on the piloting results and informal conversation with non-English major college students. (Information gathered from informal conversations with sampled students was used in the interpretation of certain piloting results.)

Items in the questionnaire were written in non-technical terms to ensure their understandability. The questionnaire contained three parts, and each part had a primary theme. Part I was about the demographic information of the participants. Part II was related to their attitudes toward the new CET 4 LCS in twelve aspects, like its difficulty, time allotment, test format, reliability and so on. Part III was about their behaviors in learning process in two perspectives: learning contents such as the quality and quantity of learning materials and learning methods which include both the coaching methods and the recommended methods. The questionnaire can be found in Appendix.

3.3 Data Collection and Analysis
Data were collected separately from 34 colleges and universities in Xi’an in November 2007 and were collected either when the students were in their English class or when they were learning in their classrooms. All the participants were given plenty of time to finish the questionnaire. As had been mentioned, the return rate was 100%, and the validity rate was 89.1%.

All the collected data were entered into the SPSS 15.0 run in Windows XP and were checked twice almost one by one. Three times of random check were conducted before operation.

4. Results and Discussion

4.1 Students’ Attitudes Toward CET 4 LCS
Students’ attitudes were operationally defined as students’ apprehensions and understandings of aspects of the new CET 4 LCS. Twelve items were examined, with the scale alpha of .7276. Results of this part are reported in Table 4.1.1.
Table 4.1.1. Students’ attitudes toward aspects of the New CET 4 LCS

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Uncertain</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Necessity of administrating the subtest</td>
<td>31.4%</td>
<td>50.5%</td>
<td>6.5%</td>
<td>7.5%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Scientificness of the subtest’s design</td>
<td>6.1%</td>
<td>33.1%</td>
<td>38.2%</td>
<td>18.8%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Conformity between subtest’s content and teaching requirements</td>
<td>3.1%</td>
<td>26.3%</td>
<td>32.4%</td>
<td>30.0%</td>
<td>8.2%</td>
</tr>
<tr>
<td>Subjective test format of the subtest</td>
<td>5.1%</td>
<td>35.2%</td>
<td>37.2%</td>
<td>19.1%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Objective test format of the subtest</td>
<td>6.1%</td>
<td>42.3%</td>
<td>34.8%</td>
<td>14.7%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Difficulty of the subtest</td>
<td>21.8%</td>
<td>35.8%</td>
<td>26.3%</td>
<td>13.3%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Validity of the subtest</td>
<td>2.4%</td>
<td>23.2%</td>
<td>21.5%</td>
<td>39.6%</td>
<td>13.3%</td>
</tr>
<tr>
<td>Adequacy of the time allotment</td>
<td>2.4%</td>
<td>26.3%</td>
<td>14.7%</td>
<td>40.6%</td>
<td>16%</td>
</tr>
<tr>
<td>Reliability of the listening subtest</td>
<td>3.1%</td>
<td>41%</td>
<td>37.2%</td>
<td>16.0%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Score reporting of the subtest</td>
<td>3.1%</td>
<td>21.8%</td>
<td>42.0%</td>
<td>28.0%</td>
<td>5.1%</td>
</tr>
<tr>
<td>Clarity of the scoring criterion</td>
<td>8.9%</td>
<td>50.2%</td>
<td>29.4%</td>
<td>7.8%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Appropriateness of the score percentage</td>
<td>9.9%</td>
<td>49.8%</td>
<td>19.5%</td>
<td>13.3%</td>
<td>7.5%</td>
</tr>
</tbody>
</table>

It can be seen from the table that students show positive attitudes toward certain aspects of the new CET 4 LCS, including the necessity of its administration, its test design, test formats (subjective and objective), reliability, scoring criterion and the appropriateness of its score percentage among the total. Aspects of the listening subtest that received negative attitudes among college students include the conformity between the subtest’s contents and teaching requirements, difficulty, validity, adequacy of its time allotment and score reporting.

It has been pointed out that the students’ attitudes toward and perceptions of a test can affect their learning process and learning product (Hughes, 1993). Hence, a positive attitude toward a certain test is a prerequisite to beneficial washback effect. In this study, students’ positive attitudes toward the above mentioned seven aspects of the LCS are encouraging in that they have indicated the necessity and success of the implementation of the new CET 4 LCS and its high quality as well. Therefore, these positive attitudes can be regarded as signs of positive washback effects.

Students’ attitudes toward a test’s difficulty and the adequacy of a test’s time allotment are directly related to their language abilities. Due to the relative small size of the research sample, this study fails to conclude students’ negative attitudes toward the two aspects as either positive or negative washback effect. However, care must be taken in future test development in terms of its difficulty and time allotment in order to measure students’ language proficiency and promote beneficial washback effects. However, this reflects a fact that the general listening ability of college students is quite low.

The researcher’s personal communication with students during the questionnaire design stage showed that college students are not familiar with the use of score report and they have no idea about its significance. The intention of using score report to replace
the CET 4 certificate was to reduce the test’s “social weight” (Gu, 2007: 1) and promote positive teaching and learning. However, the results of this study show that it fails to achieve its purpose and this is a sign of negative washback effect of the test. Purpose of using the report and its interpretation should be made clear to test-takers in order to promote beneficial washback effects of the LCS. This is the responsibility of the test designers, college English teachers and other stake-holders as well.

What is more, results show that college students generally claim that the listening subtest fails to measure their true listening comprehension ability, i.e., the listening subtest is invalid. Nevertheless, this can not be regarded as a sign of negative washback effect. In the researcher’s personal communication with students during the questionnaire design stage, they argued that they could perform better if the test format was different. It seems that the test looks invalid to them. Thus, it is hard to summarize this as either positive or negative washback effect of the test. But this result does pinpoint one problem. It is argued by some researchers that the concept of face validity should be abandoned (Zou, 2005). The results from this study seem to oppose this opinion. A good test is one that appears to measure the ability it claims to measure, as judged by an untrained observer. Results indicate that the listening subtest lacks face validity. Accordingly, face validity should be taken into consideration in future LCS development and validation process. Though the term face validity has been abandoned, its meaning can be preserved and be merged into the concept of content validity, since the “test takers will perceive tests as more interactive and authentic, and will therefore be more motivated, which could lead to enhanced preparation and hence to better performance” (Bailey, 1999: 14).

Last but not the least, the definition of “teaching requirements” in the student questionnaire was not well explained to the students, which makes it hard to interpret students’ responses to this item. This is a limitation of this study. Besides, this aspect can be the topic of another paper focusing on textbook washback, and so the present paper will not discuss this aspect in detail.

4.2 Students’ Learning Behaviors
Students’ behaviors were operationally defined as students’ actions influenced by the new CET 4 LCS in their learning process in two perspectives: learning contents and learning methods, based on suggestions of Bailey (1996), Gu (2005), Qi (2004), the new CET 4 specification and the fifteen washback hypotheses (Alderson & Wall, 1993). Fifteen items were used in this part, with the scale alpha of .7320.

Three subcategories—learning contents, coaching methods and recommended learning methods—were grouped together for the convenience of analysis and the three subcategories are reported separately in the following sections.

4.2.1 Learning Contents
Students’ learning contents influenced by the LCS were examined in six aspects, and the results of these are reported in Table 4.2.1.1.
Using textbooks and authentic materials and increasing the quantity and variety of the listening materials are preferred learning behaviors according to CECR. Therefore, they are signs of positive washback effects of the CET 4 LCS.

As for using past papers and mock tests, simple matters of positive or negative washback effects cannot be drawn without analyzing the quality of mock tests and the degrees and intensities of students’ uses. This study mainly examined the nature and scope of the listening subtest’s washback effect; therefore, the analysis of the above points is beyond the scope of this study, but it is strongly recommended as the topic of another paper.

### 4.2.2 Coaching Methods

Four aspects in students’ coaching methods influenced by the CET 4 LCS were examined in this study. Table 4.2.2.1 shows the results of the four aspects.

<table>
<thead>
<tr>
<th>Behaviors (Coaching methods)</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Uncertain</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applying test-taking strategies</td>
<td>12.3%</td>
<td>41.0%</td>
<td>18.4%</td>
<td>24.2%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Enrolling in test-preparation courses</td>
<td>10.2%</td>
<td>46.8%</td>
<td>21.8%</td>
<td>18.8%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Demanding test-preparation classes</td>
<td>17.1%</td>
<td>61.4%</td>
<td>10.6%</td>
<td>9.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Learning those that occur in the test first</td>
<td>17.4%</td>
<td>63.5%</td>
<td>9.2%</td>
<td>8.9%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

It is shown in the table that college students generally do adopt those coaching methods, and these are the negative influence the listening subtest has produced on college students’ learning. This paper defines washback as the influence of testing on learning, at a micro level. Thus, the negative influence can be regarded as a negative washback effect. However, those students’ coaching methods can never be caused by a test itself. There should be other factors, together with the test, that affect what students actually do.

### 4.2.3 Recommended Learning Methods

Students’ recommended learning methods influenced by the LCS were examined in five aspects in accordance with CECR and CET 4 syllabus. Table 4.2.3.1 objectively reports the descriptive results of those items.
Table 4.2.3.1. Students’ learning behaviors—recommended learning methods

<table>
<thead>
<tr>
<th>Behaviors (Recommended learning methods)</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Uncertain</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main idea and important details</td>
<td>24.9%</td>
<td>60.4%</td>
<td>7.2%</td>
<td>6.5%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Speaker’s attitude, opinions etc.</td>
<td>28.0%</td>
<td>64.5%</td>
<td>4.1%</td>
<td>2.4%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Implied meaning</td>
<td>29.4%</td>
<td>59.7%</td>
<td>8.2%</td>
<td>2.4%</td>
<td>.3%</td>
</tr>
<tr>
<td>Phonological features</td>
<td>25.6%</td>
<td>55.6%</td>
<td>10.9%</td>
<td>7.5%</td>
<td>.3%</td>
</tr>
<tr>
<td>Logical relationships</td>
<td>25.3%</td>
<td>61.4%</td>
<td>9.6%</td>
<td>3.4%</td>
<td>.3%</td>
</tr>
</tbody>
</table>

All the five recommended learning methods are the preferred learning methods indicated by CECR and CET 4 syllabus. As a result, they are signs of positive washback effects in this study. Accordingly, teaching and learning of those aspects should be integrated into the teaching curriculum of listening.

5. Conclusion

This study is the first attempt to investigate the washback effect of the new CET 4 LCS on language learners, which is explored through college students’ attitudes toward it and their learning processes influenced by it. While college students show their satisfactions toward the objective aspects of the listening subtest, they are subjectively unsatisfied by the subtest’s difficulty, its time allotment and its score report. Besides, both preferred learning contents and learning methods are discovered in their learning process, but coaching methods still prevail among college students. Analysis of the results shows that the new CET 4 listening subtest has produced more positive washback effects than negative washback effects on college students in China. However, this paper defines washback as the influence of testing on teaching and learning at a micro level. Results from this study indicate that a simple washback model at the micro level is far from enough to explain the complicated phenomenon of washback. Washback is a phenomenon which cannot be separated from its social and educational contexts.

Cheng (2005: 61) points out that “fewer studies have employed a mixed (quantitative and qualitative) method for the investigation of washback”, and this study also cannot overcome this problem due to some practical constraints. It is suggested that future investigation on washback to learners can use a variety of research methods as suggested by Wen (2001) to capture the nature of this issue. Besides, the size of population in this study is still not large due to some practical reasons. Since CET 4 is an exam held nation-wide, it would be more convincing if data were collected in a larger number and in more than one province. At the same time, an investigation of washback among students in middle-class universities and top-class universities will also give some insights into this issue. What is more, in view of the status quo of the listening comprehension ability of Chinese learners and the large population of college students who failed to pass the listening subtest as well as CET 4, a study on washback effect targeting toward the lower-level college students will give invaluable information about college English teaching and learning.

The existing washback literature is still insufficient in number to generate a systematic
theory (Alderson & Wall, 1993), and this study provides it with new evidence from a new perspective, i.e., learner perspective. The impact of testing on students’ motivations and performance is a thoroughly researched area in the field of general education, but surprisingly it is little explored in second language education (Chapman, 2007). Past washback studies mainly focus on language teachers while the analysis of language learners is mainly used for the interpretation of teaching activities. Since the ultimate goal of teaching and testing is to improve students’ learning, it would be valuable to examine “learner washback” (Bailey, 1996: 12). CET 4 washback studies have been carried out for some time in mainland China, and washback studies about some of its subtests have also been carried out, such as the CET-SET washback study by Jin (2000) and the CET 4 writing washback study by Chu and Gao (2006). However, there is no washback study on CET 4 LCS, since listening has always been neglected as a research subject (Shohamy & Inbar, 1991). In an earlier preliminary research carried out by this author, it was found that most non-English major college students regard listening comprehension as the skill they want to improve urgently and also the most difficult skill to improve within a short period of time. What is more, given listening comprehension’s top position in CECR, its high percentage and multiple test formats in the new CET 4, an investigation of its washback effect has significant value. In sum, this study aims at exploring the washback effect from a less explored perspective, i.e., learner perspective, together with a less explored sub-set of CET, i.e., the new CET 4 LCS, in the hope that results can make some contributions to the literature of washback in language testing.

References


Appendix　新大学英语四级听力考试对学生的反拨效应问卷调查

亲爱的同学：

我们希望了解你的英语听力学习情况和你们对将要参加的英语四级听力考试的想法，希望你能帮助我们填写此问卷。请你根据自己真实的想法和实际做法答题，你在这里所提供的信息均会得到保密。多谢您为此付出的时间。

第一部分 请选择恰当的答案打勾或填写适当的答案。

你的性别：(1) 男 (2) 女
你所在的学校是：
你所在的年级是：

4. 你以前参加过几次新四级考试？
   (1) 无 (2) 一次 (3) 两次 (4) 两次以上

第二部分 请你按照五个等级来评估以下问题。

7. 你对以下有关新四级听力考试问题的看法如何？
   1=完全同意  2=同意  3=无法确定  4=不同意  5=完全不同意
   (1) 有必要进行大学英语四级听力考试。
   (2) 四级听力考试试卷设计科学。
   (3) 四级听力考试内容与教学内容基本一致。
   (4) 四级听力考试主观题形式令人满意。
   (5) 四级听力考试客观题形式令人满意。
   (6) 四级听力考试很难。
   (7) 四级听力考试成绩能够反映自己实际英语听力水平。
   (8) 四级听力考试部分时间充裕。
   (9) 四级听力考试评分具有客观性和一致性。
   (10) 四级听力考试分数报道令人满意。
   (11) 四级听力考试评分标准明确。
   (12) 四级听力考试分数在试卷总分中比重合理。

第三部分 请你按照五个等级来评估以下问题。

12. 你对以下考试前听力学习行为的看法如何？
   1=完全同意   2=同意   3=无法确定   4=不同意   5=完全不同意
   (1) 使用教材练习听力。
   (2) 使用真实的听力材料，比如看电影，听广播等。
   (3) 使用模式试题来练习听力。
   (4) 使用真题来练习听力。
   (5) 增加听力材料的数量。
   (6) 增加听力材料的种类。
   (7) 寻求听力学习上的指导和表现上的反馈。
   (8) 应用听力考试技巧，比如不听就猜答案。
   (9) 参加听力考试辅导班。
(10) 不定期地要求听力老师在课堂上指导听力考试，为考试做准备。
(11) 逃课去准备听力考试。
(12) 首先学习那些最有可能在听力考试中出现的内容。
(13) 在听的过程中注意理解所听材料的中心思想和重要细节。
(14) 在听的过程中注意判断说话人的态度，观点等。
(15) 在听的过程中注意理解隐含的意思，如推论隐含的意思，判断话语的交际功能。
(16) 在听的过程中注意借助语言特征理解听力材料，如辨别语音特征，从连续的话语里辨别语音，理解重音或语音语调等。
(17) 在听的过程中注意理解句间关系，如比较，结果，程度，目的等。

(Copy editing: Cao Yongheng)