

A Case Study on the Washback Effects of CET-4 Listening in a Local Context

Yanli Fu

Guangdong University of Science and Technology, Dongguan, Guangdong, China

Abstract

College English Test Band Four (CET-4) which is designed according to the college English teaching syllabus, is one of the existing large-scale language tests in China. Since the inception of the test, due to the huge number of test takers, it has received a lot of attention home and abroad. During the past years, it has experienced some reforms in different aspects either wholly or partly, like scoring system, format, content, etc., which has always been the research focus. This article intends to conduct a case study on the examination of the reformed listening part of CET-4 in terms of purpose, factors affecting the test performance and assessment framework against the context of an application-oriented college, aiming at gaining some washback both on learning and teaching from teachers and students' standpoint and providing some suggestions for future teaching and learning.

Keywords

CET-4 listening; washback effects; case study.

1. Background

In order to motivate fulfilment of the national curriculum and evaluate college students' English ability, the College English Test (CET) was launched in 1987, including Band 4 (CET-4) and Band 6 (CET-6). Almost all the tertiary institutions in China require their students to take the CET-4, and those who have passed the CET-4 can choose to take the CET-6. Officially started in 1987, College English Test Band Four (CET-4) is administered by the NCETC and held twice a year, which is a general standardized high-stake test and widely recognized among Chinese higher educational institutions and employers, whose evaluation criteria (Yang & Weir, 1998) is based on the teaching syllabus and aimed at achieving a balance between linguistic knowledge and linguistic competence. In the past decade, the CET has become the most influential English test in China and also the world's largest language test administered nationwide (Jin & Yang, 2006). The design of CET-4 strictly adheres to the procedures, of which reliability and validity are the paramount indexes in terms of test quality (Yang & Weir, 1998). Since its inception, it has played a very influential role and dominated in Chinese higher education. Every college student is encouraged to sit for this test and is allowed to take it without time limit. Students who participate will get a score report accordingly. Since it is a norm-referenced test, the score report shows how well the test taker performs. What's more, with more and more attention and recognition on English both in various schools and job markets, the number of test takers are increasing exponentially. At first, passing the test was closely related with one's diploma, which means it is a prerequisite for getting an academic degree, but now it is changing and passing is not a requirement any more and is not part of official policy. It has gone through three different stages and experienced some reforms, of which is the one in 2006, a very effective testing instrument was developed with higher validity and reliability, by which students' language proficiency was evaluated. As we all know, CET-4 has gone through several modifications, some are upgrade retrofits and some are change retrofits. As for the reform of listening section in 2016, it is called upgrade retrofit. Based on

the layers of change from Brand (1994, cited by Fulcher & Davidson, 2009), it is a kind of space plan, which puts more emphasis on speaking and listening, stressing practicality.

The total score of CET-4 is 710 points with four sections (writing 15%, listening comprehension 35%, reading comprehension 35% and translation 15%) and lasts 125 minutes. In terms of listening comprehension section, it is consisted of three subsections: short news (7%), long conversations (8%) and essays (20%), which is the reformed format in 2016 and takes 25 minutes with all multiple-choices. There are three paragraphs of news listening, all of which are international current events. They are broadcast in the form of VOA and BBC. There are 2-3 questions in each article, totally 7 questions. More elements, longer sentences and loose structure are involved in the news, which increases the difficulty for the examinees to understand the news. News broadcast is different from other forms of listening materials. It has its own characteristics in structure. There are two long dialogues. After each long dialogue, there are four questions, a total of eight questions, usually in the form of a man and a woman. There is a 15 second gap between each question and the recording speed is about 120-150 words per minute. The content of the dialogue is relatively extensive, involving daily life, clothing, food, housing, transportation, study, work and other aspects we are familiar with. The sentence structure and content are not complex. There are 3 passages in the listening comprehension, and there are 3-4 questions after each passage, with a total of 10 questions, accounting for the largest proportion of listening comprehension. The short listening materials are common speeches and narratives with common themes and not too complicated plots and the words used are not beyond the scope of the CET-4 vocabulary, which mainly examines the questions of factual details, reasoning judgment and generalization.

Since the prevalence of standardized test in the 1970s, it has met challenges and harsh criticisms (Shohamy, 2007). Similarly, CET-4 is no exception, on which discussions have never ceased. The effect and reflections towards this test has received both praise and criticism as well. People who advocates (Chen & Webb, 2017) this test claim that it provides positive washback and beneficial feedback to teachers' teaching and students' learning. Keeping this test in mind, considerable emphases are placed on preparation, which give high incentives for both teachers and students. Besides, it is also regarded as one of the qualifications in relation to English proficiency evaluation for employers (Yang, 2003). By contrast, the test has also been criticized for its negative influence and inability to test students' real proficiency in English (Yu & Zhao, 2014). Since it is a very influential acknowledged high-stake test, more and more colleges are competing each other for the ranking by adopting exam-oriented policy, which make the test become more and more utilitarian (Yu, 1999). For example, instead of training students' English proficiency, "sea tactics" and "spoon-feeding method" are frequently used methods, aiming at improving CET scores, which produces students with high scores and low ability (Yu, 1999).

In terms of the purpose, overall, CET-4 is a proficiency test, from which the test taker's knowledge and skills of English are comprehensively tested and measured and an objective, accurate and impartial evaluation are presented so as to judge whether candidates have reached a certain standard. Meanwhile, prescribed teaching goals are tested as well, giving teachers and students feedback. All language tests are used to make decisions and the CET score is one of the indicators for companies and employers to recruit. It provides a scientific basis for the evaluation of teaching quality; at the same time, it also provides an important reference basis for employers in recruiting talents.

2. Literature review

The influence and the so-called “social weight” of the high-stake CET-4 has exerted teaching and learning greatly, involving different stake-holders (Gu, 2007). What’s more, the research on CET-4 is ranging from reliability, validity to washbacks (Gu, 2007). Due to the recent reform of the listening section, the significance and the purpose of the test change accordingly, it is very necessary to study the washback of this kind of change.

Theories have been proposed on the influence of testing on teaching and learning, a phenomenon commonly referred to as “washback”(Li, 2009). Studies on washback range from definitions, classifications, effects to models. Definitions regarding washback vary differently. In terms of language assessment, “washback” means that the impact it has both on teachers and students, which can be positive or negative(Liu & Gu, 2013). For different research purposes, the specific amount of sample and background, scholars may focus different aspects of the “washback”(Wang, 2010). According to Buck(1988, cited by Chen & Webb, 2017), when the test is high-stake and important, the pass rate will naturally become paramount concern, which plays a very important role in classroom teaching and can be either good or bad. Since the test is so powerful and related with decision-making, various stakeholders are involved, who will modify the curriculum in order to adapt to the the test, which can be reflected by the “washback effect” developed by Shohamy (1992). Likewise, Alderson and Wall (1993) further illustrate the effect of the washback, which motivates people to do things that they would not have done. Bachman and Pamler(1996) investigate the direct impact of washback, not only on teachers and students, but also on educational system and society. In terms of classification, Hughes(1989, cited by Chen & Webb) holds the view that if a test facilitate learning and teaching, it is positive; if it impedes learning and do something bad on the teaching or learning process, it is negative. Hence, positive and negative washback or beneficial or harmful washback are classified. In particular, Hughes further discusses the ways of promoting positive washback. Based on Hughes, Bailey(1996) also analyzed four ways of promoting beneficial washback. In the opposite, the negative washback are discussed by some scholars. For example, “distort curriculum” showed that teachers do not pay attention to activities and tasks which have no direct relation to exams (Vernon,1956, cited by Chen &Webb). What’s more, overt and covert washback also belong to types of negative washbacks. Overt washback is explicit and easy to be identified, which can be explained in the process of exam-oriented teaching and learning, that is to say, all the approaches and knowledge are test-driven. While covert washback is just the opposite, it is a kind of unconscious process and can be found in classroom activities, where not all the students are engaging(Prodromou, 1995).

The number of models on washback has a wide range. Bachman and Pamler’s(1996) pattern on washback can be viewed from micro and macro level. Micro level refers to the effect on teachers and students, while the macro level is from a more global view, that is to say, from the society or the educational system’s perspective. A very influential one is Alderson and Wall’s(1993) washback hypothesis, from which fifteen questions are presented, which Qi(2004) divides into three groups: washback on teaching, washback on learning and the degree of washback. Another model PPP developed by Hughes(1993), involves participants, process and products, which is a trichotomous division. Participants not only refer to teachers and students, but also other stakeholders. Process is any change facilitating learning and teaching. Products mean the learned knowledge and its quality. Based on Hughes, the description of Bailey’s model (1996) shows us both the direct and indirect impact on the involved parties, which is more complex. Gu(2007) developed the three “PPP” into four “PPPP”, which adds the perception, to be specific, views and opinions towards washback. Her study was carried out in both case study setting and nationwide contexts. One of her relevant findings was that the CET has motivated teachers in their teaching. In particular, she concluded that the CET writing test exerted positive washback

on college English teaching, which was manifested by the fact that students' writing scores rose rapidly since the adoption of the cut-off score policy in late 1990s (Li, 2009).

A major debate on the CET washback studies in China seem to be whether the CET brings more positive or negative influence to English teaching and learning, in this study, the author would like to study the positive washback by adopting Alderson and Wall's Wash hypothesis model and Gu's perception idea from "PPPP" model both on teachers and students in terms of learning and teaching in a specific micro context.

3. Research design

The research questions are as follows:

What factors will influence students' CET-4 listening test performance? Which are the most important ones?

Which one is the most difficult for students among the three parts in listening? How do students feel?

What aspects of listening should teacher pay more attention to in the future teaching according to students?

3.1. Methods

This study intends to adopt a qualitative method to investigate the research questions, since it is a very effective method complementary to quantitative method and is a trend in applied linguistics (Read, 2010), which is appropriate for solving the research questions. Interviews can provide important insights to the studied questions (Yin, 2018), and accordingly semi-structured interviews are used as a method to collect data in this study.

3.2. Participants

Six non-English major sophomores are invited with their approval to take part in the interview, who come from different major background, and all of them have experienced the CET-4 test. Before the interview, an interview schedule and protocol are designed. At the same time, all the ethical issues are properly dealt with. The interview is semi-structured, which allows students to express their views freely and thoroughly. Each interview takes around thirty minutes, considering the number of participants, totally more than three hours. All the interviews are conducted individually in quiet places on campus. Regarding all the students are Chinese, in order to avoid misunderstandings and make students behave naturally and elaborate in great detail, the interview is conducted in Chinese. Meanwhile, all the interviews are recorded by a recording pen, which is used for later data analysis. During the interview process, the author keeps taking notes and memos, which can also be used to analyze data.

3.3. Data analysis

In accordance with the research questions, the interview questions are designed. By employing the six steps of thematic analysis (Braun & Clarke, 2006), the interview transcripts and research notes are analyzed. Firstly, the oral interviews are transcribed into printed text. As for the research question one, almost all of them believe listening is the most challenging part in CET-4, and the mastery of words, concentration, motivation, classlessness, anxiety, test skills and time are the factors influencing their listening performance. With regard to the most important factors, all of them mentioned the test skills and vocabularies. Among the three parts in the listening test, all of them think that the long conversation is the easiest to understand, while five believe the passage is the most difficult one and one student thinks both the news and the passages are challenging. As for the suggestions for the future teaching, they hold the view that the interpretation of passages should receive more focus.

4. Discussions

4.1. Washback effects on teachers

On the basis of the analyzed data, some washback effects on teachers, especially on the future teaching contents and approaches are identified. Generally speaking, the washback is positive, which can facilitate future teaching. In terms of the most important factors affecting students' listening performance, teachers should deliberately train students to master some test management strategies and wiseness strategies according to their English proficiency levels, taking their individual differences into consideration. In addition, teachers can purposely practice students' listening to reduce their anxiety. Instead of just employing product-oriented approach, process-oriented method is strongly encouraged. In view of students' perception towards the easiest part, communicative-based teaching should be adopted in the classroom teaching, which is more understandable and effective, because students have the advantage of using this approach based on the data. As for the future teaching contents and materials in relation to listening, passages or monologues, sometimes news should be put on the first priority according to the interview data. The ways to master new words should also be instructed and guided, after all, some of them feedback that vocabulary is one of factors influencing their listening performance.

Regarding the goal of college English teaching is to cultivate students' comprehensive English application ability, especially listening and speaking ability, in the future classroom teaching, teachers need to consciously train students listening instead of ignoring it.

4.2. Washback effects on students

Speaking of the beneficial washback, it may provide some directions for students' future learning and preparation for CET-4. As far as the students are concerned, they can keep practicing listening within the time limit to overcome anxiety, carelessness and become more concentrated. After all, practice makes perfect. Almost everyone believes that vocabulary is a decisive factor with respect to their listening performance. Therefore, for future learning, students should spend more time on words, since they are the foundation of listening, if they do not know the word, how could they understand the listening, let alone the passages? Almost all of them are likely to gain correct answers from long conversations comparing with other parts, so they need to pay more attention to passages and spend time in practicing exercises. In order to better cooperate with teachers, students are highly recommended to preview the contents and materials in advance.

5. Conclusion

This study makes a simple research on the washback effect of CET-4 from the micro perspective of students and teachers, which can be regarded as a pilot study. The results show that the new listening test format brings positive washback on both teachers and students, promotes the implementation of the requirements on improving students' communicative competence in the College English Syllabus, and stimulates students' interest in listening. Although students say that the part of listening test is challenging, they all hold a positive attitude towards listening learning. To some extent, it reflects that the reform of listening test has a great impact on English listening learning attitude, the depth and breadth of the learning content. The increasing listening score makes students pay more attention to listening, spend more time on listening practice and improve their listening ability. The washback effect of CET-4 is very complex, involving all levels and various factors. Based on this, we should formulate the specific steps of the washback effect study, clarify the specific tasks of each stage, making the research more scientific and comprehensive, extending the positive washback effect on the test to a

greater extent, improving students' language communicative ability and language application ability so as to realize the ultimate goal of foreign language teaching reform.

References

- [1] Alderson, C. & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14 (2), 115-129.
- [2] Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford University Press.
- [3] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101.
- [4] Chen, X., & Webb, J. G. (2017). The CET-4: Encouragement or Impediment of Undergraduate English Teaching and Learning. *International Journal of Education and Human Developments*, 3(4).
- [5] Fulcher, Glenn, & Davidson, Fred. (2009). Test architecture, test retrofit. *Language Testing*, 26(1), 123-144.
- [6] Gu, X.D. (2007). An Empirical Study of CET Washback on College English Teaching and Learning in China [J]. *Journal of Chongqing University (Social Science Edition)*, 4.
- [7] Jin, Y. & Yang, H. (2006). The English proficiency of college and university students in China: as reflected in the CET. *Language Culture and Curriculum*, 19(1), 21-36
- [8] Li, H. (2009). Are teachers teaching to the test? A case study of the College English Test (CET) in China. *International Journal of Pedagogies & Learning*, 5(1), 25-36.
- [9] Liu and Gu. (2013). A review of the international and national empirical researches on washback in the past 20 years. *Foreign Language Testing and Teaching*(1), 6-19.
- [10] Prodromou, L. (1995). The backwash effect: from testing to teaching. *ELT Journal*, 49(1), 13-25.
- [11] Qi, L.X. (2004). *The Intended Washback Effect of the National Matriculation English Test in China: Intentions and Reality*. Beijing: Foreign Language Teaching and Research Press.
- [12] Read, J. (2010). *Researching Language Testing and Assessment*. Continuum companion to research methods in applied linguistics, 286.
- [13] Shohamy, E. (1992). Beyond proficiency testing: a diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76(4), 513-521.
- [14] Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education: Principles Policy & Practice*, 14(1), 117-130.
- [15] Wang, W.L. (2010). Investigating the Washback Effect of the New CET 4 Listening Comprehension Subtest on Language Learners. *Chinese Journal of Applied Linguistics (Foreign Language Teaching & Research Press)*, 33(5).
- [16] Yang, H. & Weir, C. (1998). *Validation study of the national college English test*. Shanghai: Shanghai Foreign Language Education Press.
- [17] Yang. (2003). The 15 years of the CET and Its Impact on Teaching. *Journal of Foreign Languages* (3), 21-29.
- [18] Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Los Angeles, CA: SAGE.
- [19] Yu, H.R. (1999). What kinds of personnel of foreign languages are needed in the 21st century? An issue beginning with the negative effect of CET-4 and CET-6. *Journal of Jinling Polytechnic College (Social Science Edition)* (4), 33-36.
- [20] Yu, Qinglan, & Zhao, Yushan. (2014). Moving forward to Communicative Proficiency: A Comparison between CET-4 Listening Comprehension Tests and IELTS Listening Subtest. *Journal of Language Teaching and Research*, 5(6), 1322.