Investigating the Relationships Between Listening Skills and Genre Competence Through Cognitive Diagnosis Approach

Huilin Chen¹ and Jinsong Chen²

Abstract

Although research on listening skills has been frequently conducted to discover the nature of listening comprehension, there is little study about listening genre competence which is related with knowledge about listening text types. In order to find out whether listening skills and listening genre competence are related, cognitive diagnosis, a quantitative method to disclose finer-grained latent attributes, was adopted in this study. The generalized deterministic inputs, noisy “and” gate (G-DINA) model, which takes attribute compensation and attribute interaction into consideration, was used to carry out cognitive diagnostic analysis. The listening comprehension subtest of Band 4 of Test for English Majors (TEM) which is a large scale English proficiency test for English Majors in China was used as the proficiency test for homogenizing the participants. Three genres in the subtest, dialog, lecture, and news, were investigated. The 2,285 subjects were sophomore English major college students and also test-takers of the same TEM4 examination. They were chosen by random sampling from the nationwide test population in China. The study analyzed three types of relationships between listening skills and genre competence. By analyzing how mastery of certain listening text genres goes with mastery of listening skills according to latent class distribution, the coexistence relationship was discovered. By comparing the average number of skills/genres mastered when the number of genres/skills mastered increases through One-Way ANOVA, compensatory and contributory relationships were revealed. The study also found that the subjects mastering Lecture genre got higher listening scores.

Keywords
cognitive diagnosis, listening skill, listening genre competence, coexistence, compensatory, contributory

Introduction

In order to communicate successfully in different situations, English language learners need to listen to texts of various types. Texts from the same type usually have similarities in form, style, and structure. They are organized in the same format and are used for the same social purpose. The types of spoken and written discourse recognized by a discourse community are termed as genres and each genre has some typical features: linguistic (particularly grammatical or lexical choices), paralinguistic (e.g., print size, gesture), or some contextual and pragmatic (e.g., setting, purpose) (Johnson & Johnson, 1999).

Genre competence was also termed as discourse competence (Connor & Mbaye, 2002) or discoursal knowledge (Weir, 2005), and related to discourse construction which was one of the listening processes in the cognitive validity for listening (Field, 2013). Genre competence in this study refers to learners’ familiarity of certain genres and their ability of understanding and structuring texts of certain genres. There is a close relationship between genre competence and text understanding, but there was little research on the relationship between listening genre competence and listening text understanding. Therefore, this paper attempts to use a cognitive diagnostic method which is sensitive to the estimation of finer-grained latent variables to explore the possible relationship between listening skills and listening genre competence, so as to further explore the process and nature of listening comprehension.

Research Background

Genre and Genre Competence

The term “genre” once was mainly used to classify literary works. Now genres are not limited to the field of literature. From the 1970s, linguists began to introduce the concept of...
Genre competence in linguistic studies. There are mainly three approaches to linguistic studies of genres (Hyon, 1996). The New Rhetoric approach (Bazerman, 2009; Freedman & Medway, 1994; Miller, 1984) focused more on the situational contexts than on their forms and defined genre as a social action. The ESP approach (Bhatia, 1993; Swales, 1990) regarded genres as oral and written text types defined by their formal properties as well as by their communicative purposes. The Australian systemic functional approach, represented by Martin (1985) and Nunan (2008), perceived genres as staged, purposeful, socially-constructed communicative events which generally result in spoken and written texts. Based on the above genre theories, the concept of genre has the following features: (1) Genres are determined by a clear communicative purpose. (2) Genres exist in a set of prescribed rules. (3) Genres exist in both written and spoken texts.

Listening texts possess all the above features of genres. Different listening text types, like dialogs, lectures, news, and so on, serve quite different communicative and social purposes. Listening text types are characteristic of special formal properties which are usually related with formality and wording. Listening texts belong to spoken texts and therefore are possible vehicles conveying genres.

Since genres are related with the features of texts, genre competence can be defined as learners’ familiarity of text features of genres and their ability of understanding and structuring texts of genres. Language competence concerning text features has long been included as an important component in major models of language ability. There is a discourse competence component in the model of communicative competence developed by Canale and Swain (1980) and Canale (1983) and a textual competence subcomponent in the model of communicative language ability proposed by Bachman (1990). Connor and Mbaye (2002) further defined discourse competence as the knowledge of the discourse organization of genres in addition to cohesion and coherence. Based on the cognitive validity model (Weir, 2005), Field (2013) emphasized the importance of discoursal knowledge in listening processes. Genre competence in this study only focuses on knowledge of text types while discourse competence or textual competence is also concerned with knowledge of cohesion and coherence. In a word, genre competence belongs to discourse or textual competence and is indispensable in language ability and also in listening processes.

**Genre Competence in Listening Comprehension**

In recent years, more and more research tried to find out the role of genre in language teaching. Most of the research results have proved the important influence of genre on reading (Francis & Hallam, 2000; Littlefair, 1991; Toledo, 2005) and writing (Flowerdew, 2000; Henry & Roseberry, 1998; Myskow & Gordon, 2010), but there are fewer studies in the field of listening. Shohamy and Inbar (1991) held that listening comprehension tests should include a number of texts which encompass a sample of a range of genres on the oral/literate continuum so as to ensure sound construct validity of listening comprehension tests. Berne (1992) found that although genre was not a significant factor affecting the understanding of the theme of the text, it could significantly affect the understanding of the details. A series of studies showed that many English learners who had sufficient English language ability but lacked competence of the lecture genre failed to grasp the main points of lectures (Allison & Tauroza, 1995; Thompson, 1994; Young, 1995). Manzouri (2015) found that genre teaching in listening was effective for both high-level and low-level students. Unlike the above studies, Nushi and Orouji (2020) did not include genre or text type as an important source of difficulty in listening comprehension and Metruk (2019), in a small sample study, did not find significant impact of English movies and TV programs on listening skills of EFL learners.

However, the above studies only focused on whether there was overall impact of genre competence on listening comprehension, and failed to further explore the relationship between genre competence and listening skills. In other words, previous research paid more attention to the possible role of genre competence in promoting listening performance but neglected that the underlying mechanism for listening performance may be realized by proper interplay between genre competence and listening skills.

**Listening Comprehension Process**

According to Vandergrift (2007), listening comprehension is a compensatory process. When learners are hindered in listening comprehension due to lack of language knowledge, they can rely on all kinds of compensatory information, including pictures, videos, text notes, cultural information, encyclopedia knowledge, common sense of life, and other relevant clues that can be extracted from their own minds. Listening comprehension generally involves the compensation of cognitive processes at different levels. The higher level process (top-down) begins with the understanding of the characteristics of the text and the nature of the world. The lower level process (bottom-up) starts from interpreting speech and determining meaning through sounds, words, and phrases. The top-down and bottom-up processes interact with each other and the lack of information at one level can be compensated by information at the other level.

Research in the field of educational psychology has much interest in the schema of listening comprehension. According to Peterson (1991), speech learning is easier when information is divided into meaningful blocks linked with existing meaning structures in the brain. McLeod and McLaughlin (1986) found that as learners became more and more proficient, they could reorganize the language input into a larger language pattern and understand it as a whole (as cited in
Peterson, 1991). People call these meaning structures in the brain “framework,” “script,” or “schema.” According to Carrell and Eisterhold (1983), there are two kinds of background information: content schema and formal schema. Content schemata include cultural knowledge, topic familiarity, and previous experience in a certain field. Formal schemata are related to people’s knowledge of discourse forms: text types, rhetorical conventions, and discourse organization. Both content schema and formal schema can help the listener to understand the text (as cited in Peterson, 1991).

Since listening skills mainly involve interpreting speech and understanding meaning in lower cognitive process while listening text genre competence is related with background knowledge and formal schemata in higher cognitive process, there may exist bottom-up or top-down relationships between listening skills and text genre competence.

**Cognitive Diagnosis and Listening Test**

The purpose of cognitive diagnostic assessment is to measure/evaluate an individual’s specific knowledge structure and processing skills (Gierl et al., 2005). Cognitive diagnosis infers the unobservable knowledge state of an individual according to the observable response patterns obtained from tests. Cognitive diagnosis is a substantial evaluation method aimed at exploring the cognitive process. It can find students’ differences in internal cognitive process or knowledge structure so as to provide more information for individualized teaching and teaching remedy. By fully integrating the research paradigms of cognitive psychology with recent achievements on internal mechanism of cognitive processing, psychological measurement models for cognitive diagnosis purposes, in other words, cognitive diagnosis models (CDMs for short), were developed.

Recently, cognitive diagnosis approach has also been applied to the study of listening comprehension, but the number of studies is small. Sawaki et al. (2009) reported in a CDM research that the latent ability of a listening test could be explained by three dominant skills: understanding general and specific information, understanding text structure and speaker’s intention, and understanding the connection of ideas. Yi (2017) found that in general, compensatory model, where mastery of some attributes may compensate for non-mastery of others, could better fit listening data. By adopting cognitive diagnosis approach, Min and Xiong (2019) discovered that strategies on solving problems in listening tests can compensate for the lack of the competence of speech recognition.

A special advantage of CDMs is that they can analyze multiple attributes represented in one item. Therefore, in a CDM approach, analyzing listening skills alongside with text genre competence can be realized and how listening skills and text genre competence interplay with each other can also be investigated.

**Research Questions**

According to the literature review, genre competence involves a higher level of cognitive process. Understanding listening texts from understanding listening genres should be a top-down process, while understanding listening texts from the perspective of listening skills can be regarded as a bottom-up process. How the two processes interplay with each other is worth investigating. Moreover, there is limited empirical research on the feasibility to include listening genre competences as attributes in CDM analysis and on the relationship between listening genre competence and listening test performance. Therefore, the following research questions are proposed.

1. How can analyzing listening skills alongside with listening genre competence be realized through a CDM approach?
2. What are the relationships between listening skills and listening genre competence?
3. How is listening genre competence related with listening test performance?

This research will explore the relationship between listening skills and genre competence by using the method of cognitive diagnosis, find out the relationship patterns, and put forward proposals that are helpful to optimize teaching and testing of listening comprehension.

**Materials and Methods**

**TEM4 and the Listening Subtest**

The research adopted the listening comprehension subtest of a Band 4 Test for English Majors (TEM4) as the research material. The Test for English Majors is a large-scale test designed to assess English language proficiency of undergraduate English majors in China. It is used to determine whether examinees have met the learning requirements specified in the teaching syllabus, and provide teachers with feedback on their teaching effectiveness and students on their strengths and weaknesses in English learning.

There are two levels of the test: the Band 4 is required to be taken by all sophomore English major students, while the Band 8 is an optional test administered to senior English major students. TEM4 and TEM8 are separate tests administered annually to students at the two stages of university study. TEM4 and TEM8 are comparable to First Certificate in English (FCE) examination and Certificate in Advanced English (CAE) examination respectively because they are roughly at B2 and C1 levels of CEFR (Common European Framework of Reference for Languages) respectively (Liu & Wu, 2019; North, 2007; Yang & Liu, 2019).

The listening subtest of TEM4 includes three sections: dialog listening, lecture listening, and news listening.
items for each section, 30 items in total, all of which are dichotomously scored items. The dialog listening is concerned with everyday conversation full of repetitions, redundancies, interruptions, pauses, simple, and unfinished sentences. The lecture listening is in the form of a monolog based on written notes delivered for the purpose of describing academic topics. The news listening is a rewritten edited monolog, delivered for the purpose of reporting social and political events. Although TEM8 also has a listening subtest, more focus is laid on reading, writing, and translation and the number of listening items is much smaller.

**Subjects**

The research data were provided by the Board of TEM Examinations. Since TEM4 is only administered to sophomore English major college students in China, the subjects of this study are sophomore English major college students and also test-takers of the same TEM4 examination. They were randomly sampled from all the 236,586 test-takers of the test nationwide with a selection proportion set at about 1% by using the SPSS case selection function. Random sampling was conducted because too large sample size (over 10,000) may not be suitable for CDM analysis (Based on email communication with CDM developers). With randomization, each member of the population has an exactly equal chance of being selected and the subjects selected could well represent the whole test population.

**Instruments and Procedures**

The CDM analysis was carried out with the G-DINA analysis program based on OX Edit software (de la Torre, 2011). Relationships between listening skills and genre competence were analyzed with SPSS.

In this study, the process of CDM analysis involves skill definition, Q-matrix construction, model fit, distribution of latent classes, and attribute mastery profile of each individual, which will be elaborated about in the CDM analysis section.

An important output of the CDM analysis is the table displaying distribution of latent classes. Each latent class represents a skill profile, and is characterized by a vector of 1 and 0 corresponding to the K skills, where 0 and 1 indicate the nonmastery and mastery of the skills. By analyzing how mastery of certain listening text genres goes with mastery of listening skills, the coexistence relationship between listening skills and genre competence can be discovered.

Another output of CDM analysis is attribute mastery profile for each subject. Based on the profiles of all subjects, the average number of listening skills mastered by the subjects mastering a certain number of genres can be calculated. In the same way, the average number of genres mastered by the subjects mastering a certain number of listening skills can also be calculated. If the average number of listening skills mastered increases as the average number of genres mastered decreases, mastery of listening skills compensates for mastery of genres. If the average number of genres mastered increases as the average number of listening skills mastered decreases, mastery of genres compensates for mastery of listening skills. If the average number of listening skills mastered and the average number of genres mastered increase at the same time, mastery of listening skills and mastery of genres contribute to each other. By integrating with the listening test scores, the influence of compensatory and contributory relationships on listening test performance can also be revealed. All the comparisons were conducted through One-Way ANOVA.

**CDM Analysis**

**Model Selection**

The generalized deterministic inputs, noisy and gate (G-DINA; de la Torre, 2011) model is a saturated CDM. Its formula is as follows:

\[
P\left(\alpha_{jk}^*\right) = \delta_{j0} + \sum_{k=1}^{k'} \delta_{jk} \alpha_{ik}' + \sum_{k'=1}^{k'1} \delta_{jkk'k} \prod_{k=1}^{k-1} \alpha_{ik'} - \delta_{kk'1} \prod_{k=1}^{k-1} \alpha_{ik'}
\]

In the formula, \(j\) represents a certain item, \(k\) represents the number of cognitive skills measured by the item, \(P\left(\alpha_{jk}^*\right)\) represents the main effect of specific skills and the sum of their interactions, \(\delta_{j0}\) is the guessing effect, \(\delta_{jk}\) refers to the main effect of mastering a skill \(\alpha_{jk}\), \(\delta_{jkk'k}\) refers to the interaction effect of skills \(\alpha_{ik}\) and \(\alpha_{ik'}\), \(\delta_{kk'1}\) refers to the interaction effect of skills \(\alpha_{ik}\) to \(\alpha_{ik'}\). The formula above covers all possible item effects (e.g., intercept or guessing, main, and interaction effects) and thus represents the saturated model. With appropriate constraints, different reduced or simpler CDMs can be derived from the saturated model.

The reduced models only consider some of the above item effects (e.g., only the guessing and final interaction effects are considered in the DINA model) which can be further classified into non-compensatory models and compensatory models. Under a non-compensatory model, an item can be successfully answered only if all the required attributes have been successfully mastered. In contrast, under a compensatory CDM, successfully mastering only some of the attributes may compensate for nonmastery of others (Chen & Chen, 2016). The saturated model is different from other compensatory models mainly because the former covers all possible attribute interactions.

Due to the interaction and compensation processes in listening comprehension, it is more suitable to adopt the saturated model which is not only a compensatory model but
includes all attribute interactions. Therefore, this study adopted the saturated G-DINA model.

**Skill Definition**

Through the process of defining skills, the listening skills involved in the test can be revealed. According to the test specifications of TEM4 (2000), the teaching requirements for listening for sophomore English majors are “to understand the conversation about daily life and social life of the people in English speaking countries; to understand the listening materials of medium difficulty (such as the passage in TOEFL), to understand the general idea, to understand author’s attitudes, feelings, and real intentions. Be able to understand the main content of VOA and BBC news programs at the normal speed. Be able to identify English varieties (such as American English, British English, and Australian English).”

Based on the specifications of TEM4 and well-known listening skill taxonomies (Goh & Aryadoust, 2015; Richards, 1983; Valette, 1977), the researchers combined similar listening skills and summarized the following list of listening comprehension skills for experts to analyze the test items and code the skills assessed in each item.

According to the skill definitions listed in Table 1, five English teachers and two graduate students engaged in English teaching coded the listening comprehension skills of TEM4 independently. According to the skill coding results reached by those teachers and postgraduate students, an inter-rater agreement survey on each skill coding of each item was conducted. If more than half (four or more) of the above seven raters agree with a certain skill coding of an item, then the skill coding is recognized as a listening comprehension skill of the item. If fewer than half agree, then the skill coding is regarded as invalid. According to the inter-rater agreement on listening comprehension skills among the seven raters, the TEM4 listening comprehension items used in this study possessed five skills. The other five skills, that is, Distinguishing speech, Identifying English variants, Understanding lexical and grammatical features, Identifying intention, attitude and rhetorical purpose, and Evaluating texts, failed to reach majority in the inter-rater agreement survey and were not included in the Q-matrix.

To the selected five listening skills were added the three attributes about the understanding of the three listening genres consistent with the three listening sections in TEM4, that is, dialog listening, lecture listening, and news listening so as to construct a Q matrix with eight attributes. The three attributes about listening genre competence were named as Dialog genre, Lecture genre, and News genre respectively for the sake of conciseness. Through the coding process, the researchers arrived at an initial Q-matrix which served as the basis of the Q-matrix validation process.

**Q-Matrix Validation**

The construction of the Q-matrix was realized through a series of Q-matrix validation estimations with the G-DINA analysis program based on OX Edit software (de la Torre, 2011). In order to detect the item-level misspecifications, the researchers adopted two statistics: the residual between the observed and predicted correlation of item pair with the Fisher transformation (the \( r \) statistics) and the residual between the observed and predicted log-odds ratios of pair-wise item responses (the \( l \) statistics). By averaging all \( z \)-scores of residuals related to specific item, \( sr_j \) and \( sl_j \) were

### Table 1. Definition of Listening Skills.

| No. | Skill                                      | Definition                                                                 |
|-----|--------------------------------------------|---------------------------------------------------------------------------|
| 1   | Distinguishing speech                      | Be able to distinguish the phonetic and intonation features that affect listening comprehension |
| 2   | Identifying English variants               | Be able to identify English dialects and variants that affect listening comprehension |
| 3   | Understanding lexical and grammatical features | Be able to understand the lexical and grammatical features that affect listening comprehension |
| 4   | Retrieving explicit information            | Be able to search and retrieve explicit information at a single location.  |
| 5   | Comparing similar information              | Be able to identify and compare similar explicit information at multiple locations.  |
| 6   | Understanding paraphrasing                 | Be able to understand other expressions of explicit information.           |
| 7   | Summarizing main ideas                     | Be able to summarize the main content and central ideas of chapters and paragraphs.    |
| 8   | Identifying intention, attitude and rhetorical purpose | Be able to perceive author’s intention, attitude and rhetorical purpose. |
| 9   | Making inferences                          | Be able to infer implicit content information.                             |
| 10  | Evaluating texts                           | Be able to evaluate main purpose and significance of listening materials. |

*Sources. Goh and Aryadoust (2015), Richards (1983), test specifications of TEM4 (English Major Division of National Foreign Languages Advisory Board, 2000), and Valette (1977).*
obtained. The item with the maximum values for $sr_j$ or $sl_j$ was considered as most likely misspecified. Modification was made to the coding of the most likely misspecified item in each validation estimation with the help of the above experts. After each modification, the test-level absolute model fit of the CDM analysis was updated. Chen (2017) elaborated on the details of the process. The process continued until the model fit was achieved.

Still based on the $r$ statistics or the $l$ statistics, the absolute model fit at the test level is demonstrated by the null hypothesis that the residuals are statistically insignificant (i.e., approach 0). Based on the standard error of the residuals, their $z$-scores and associated $p$-values can be calculated, and accordingly the absolute model fit evaluation at certain significance level (Chen et al., 2013) can be conducted.

Based on the initial coding conducted in the skill definition process, a series of validation estimations were conducted followed by corresponding modifications. The process of Q-matrix validation is shown in Figure 1.

The breakthrough modifications occurred to Item 8 and Item 11. After Item 8 was identified as the most likely misspecified item, the seven experts analyzed the script text and the item shown below.

**Why didn’t Bill want one of them?**
A. He wanted to buy one from Japan.
B. He wasn’t sure about its quality.
C. He thought it was for business use.
D. He thought it was expensive.

They found that the item involved information which was not directly stated in the listening script and reached a consensus to add the skill Comparing similar information to the item. After the modification, the test-level absolute model fit was updated and reached the final significance level at .01.

In the following validation estimation, Item 11 was identified as the most likely misspecified item, the seven experts again analyzed the script text and the item shown below.

**In the old days dogs were used for the following EXCEPT**
A. Hunting other animals.
B. Driving sheep.
C. Guarding chickens.
D. Keeping thieves away.

They found that the listening script contained several pieces of information which may be related with the use of dogs and reached a consensus to add the skill Comparing similar information to the item. After the modification, the test-level absolute model fit was updated again and reached the final significance level at .05. The researchers continued the validation process but could not make further improvement about the model fit based on agreed modifications. Table 2 summarizes the above modifications and the subsequent updates of test-level absolute model fit.

The model fit can be deemed satisfactory if the maximum $z$-scores are less than the Bonferroni adjusted critical $z$-score $z_c$ at a particular significance level (Chen et al., 2013). A good model fit ensures the accurate results of CDM analysis. Table 3 below displays the final Q matrix.

**Results**

**Coexistence Relationship Between Skills and Genre Competence**

One of the main advantages of cognitive diagnosis is that it can estimate the distribution of the latent classes, which can reflect different attribute coexistence patterns in the sample group. The larger proportion a latent class accounts for in the sample group, the more possible the attributes in that latent class can coexist.

In a latent class, the number order is based on the order of the attributes listed in the Q-matrix. That is, the number in the latent classes, from left to right, represent Retrieving explicit information, Comparing similar information, Understanding paraphrasing, Summarizing main ideas, Making inferences, Dialog genre, Lecture genre, and News genre. The first five attributes are listening skills, and the last three attributes are about listening genre competence. Each latent classes listed in the above table accounted for more than 1% in the sample group, and the latent class whose proportion was less than 1% was not listed.

In Table 4, subjects who mastered the Lecture genre (010), the Dialog genre and Lecture genre (110), the Lecture and News genre (011), and all three genres (111) mastered...
more listening skill combinations and accounted for larger proportions of the sample (23.95% for 010, 13.77% for 110, 16.21% for 011, and 27.65% for 111), which shows that there is a skill-genre coexistence relationship in the above latent classes.

In other latent classes, the coexistence relationship between listening genre competence and listening skill is relatively weak. The Dialog genre (100) could coexist with the two skills: Comparing similar information and Summarizing main ideas, but the proportion was only 1.71%. The News genre (001) could not easily coexist with a listening skill since no such latent classes accounted for more than 1%. In Table 4, there was no latent class including Dialog genre, News genre and listening skills at the same time, which not only demonstrates that the combination of Dialog genre and News genre is not easy to coexist with listening skills but also shows that the two genre attributes themselves are not easy to coexist.

### Table 2. Modifications and Model Fit Updates.

| Most likely misspecified Item | Modification | Updated max z of r | Updated max z of l |
|------------------------------|--------------|--------------------|--------------------|
| 8 01000100 → 01001100 | 4.1786 | 3.9058 |
| 11 10000010 → 11000010 | 3.8191 | 3.8424 |

\[ z_c \text{ score (Bonferroni correction):} \]

- \( p = .10 \) 3.6837 3.6837
- \( p = .05 \) 3.8567 3.8567
- \( p = .01 \) 4.2337 4.2337

### Table 3. Q-matrix.

| Item no. | Retrieving explicit information | Comparing similar information | Understanding paraphrasing | Summarizing main ideas | Making inferences | Dialog genre | Lecture genre | News genre |
|----------|--------------------------------|-------------------------------|---------------------------|------------------------|------------------|--------------|--------------|-----------|
| 1        | 1                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 2        | 1                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 3        | 1                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 4        | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 5        | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 6        | 1                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 7        | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 8        | 0                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 9        | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 10       | 0                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 11       | 1                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 12       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 13       | 0                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 14       | 0                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 15       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 16       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 17       | 0                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 18       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 19       | 0                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 20       | 0                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 21       | 0                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 22       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 23       | 0                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 24       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 25       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 26       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 27       | 0                               | 0                             | 0                         | 1                      | 0                | 1            | 0            | 0         |
| 28       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 29       | 1                               | 0                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
| 30       | 0                               | 1                             | 0                         | 0                      | 0                | 1            | 0            | 0         |
The subjects who mastered no genres (000) could at the same time master the four listening skills: Retrieving explicit information, Comparing similar information, Understanding paraphrasing, and Making inferences, but the proportion was only 1.71%, which shows that a small proportion of subjects who do not master any listening genres can master more listening skills than expected.

### Compensatory and Contributory Relationships

Cognitive diagnosis can not only estimate the distribution of latent classes of the sample group but also produce the attribute mastery profile of each individual which contains the subject’s mastery probability of each attribute. The mastery probability over .5 can be regarded as mastering the attribute. Otherwise, it is regarded as not mastering the attribute. By investigating the attribute mastery probability of each student together with his/her listening test score, the researchers can not only find other relationships between listening genre competence and listening skills but also find the influence of the relationships on listening test performance.

If the coexistence relationship exhibits whether listening skills and listening genre competence are in harmony or in conflict, compensatory and contributory relationships then can exhibit how listening skills and listening genre competence may help each other. Compensatory relationship between listening skills and listening genre competence, in this study, refers to the notion that mastery of listening skills can compensate for lack of listening genre competence and mastery of listening genres can compensate for lack of listening skills. Contributory relationship between listening skills and listening genre competence, in this study, refers to the notion that mastery of listening skills can promote mastery of listening genres, or vice versa.

According to the definition of compensatory relationship, more listening skills may be mastered than expected when there is a lack of listening genre mastery and more listening genres may be mastered than expected when there is a lack of listening skill mastery. In order to find out what the fact is, the researchers compared the average numbers of the skills and genres mastered at different situations. Figure 2 not only displays the average numbers of the listening skills mastered when different numbers of listening genres were mastered but also displays the corresponding average test scores.

Figure 2 shows the change of average listening skill numbers as the number of listening genres increases, accompanied with the change of corresponding test score means. The bars in Figure 2 represent the average numbers of listening skills mastered at different listening genre situations. When no listening genre was mastered, the average number of listening skills mastered was 3.2469 which was just above the moderate level. The average number of listening skills mastered increased slowly and then decreased when the number of listening genres mastered increased from 0 to 2, but increased greatly when the number of listening genres mastered increased from 2 to 3. The subjects mastering no listening genre even mastered more listening skills than those mastering two listening genres. The line in Figure 2 represents the change of test score means when the number of listening genres mastered increased. When no listening genre was mastered, the test score mean was 15.6790 which was just above the moderate level. The test score means increased slowly when the number of listening genres mastered increased from 0 to 2, but increased very fast when the number of listening genres mastered increased from 2 to 3.

In order to find out whether the change of average number of listening skills mastered and the change of test score means are significant or not, the researchers conducted One-Way ANOVA analyses to compare average numbers of listening skills mastered and test score means respectively at different listening genre mastery situations. Table 5 was derived from the multiple comparison results of the One-Way ANOVA analyses.

The multiple comparison results of the One-Way ANOVA analyses were based on the Games-Howell method because group sizes were quite different and equal variances were not assumed. The difference was considered significant when p < .05.

### Table 4. Distribution of latent classes.

| Latent class | Skill | Genre competence | Proportion (%) |
|--------------|-------|-------------------|---------------|
| 11111        | 111   |                   | 22.66         |
| 00001        | 011   |                   | 10.81         |
| 10111        | 110   |                   | 8.51          |
| 11111        | 010   |                   | 7.50          |
| 11110        | 111   |                   | 3.93          |
| 10110        | 011   |                   | 3.35          |
| 10110        | 110   |                   | 2.72          |
| 11110        | 011   |                   | 2.40          |
| 10110        | 010   |                   | 2.18          |
| 11001        | 010   |                   | 1.79          |
| 11001        | 011   |                   | 1.79          |
| 11101        | 000   |                   | 1.71          |
| 01010        | 100   |                   | 1.71          |
| 10001        | 010   |                   | 1.65          |
| 11101        | 010   |                   | 1.60          |
| 11110        | 110   |                   | 1.47          |
| 10100        | 010   |                   | 1.41          |
| 11110        | 010   |                   | 1.27          |
| 11110        | 011   |                   | 1.21          |
| 11101        | 010   |                   | 1.12          |
| 00100        | 010   |                   | 1.08          |
| 10000        | 010   |                   | 1.07          |
| 10100        | 111   |                   | 1.06          |
| 11010        | 010   |                   | 1.00          |
Table 5 reflects that there may exist either compensatory or contributory relationships between listening skills and listening genre competence. The average number of listening skills mastered and the test score means were both above the moderate level when no listening genre was mastered, which demonstrates that mastering a moderate number of listening skills may compensate for the lack of listening genre mastery and play an important role in maintaining test score means above the moderate level. When the number of genres mastered increased from 0 to 1, the average number of listening skills mastered increased but not significantly. At the same time, the test score means increased significantly. The phenomenon demonstrates that compensations from listening skills were still important when only one listening genre was mastered. The average number of listening skills mastered decreased significantly but the test score means still increased significantly when the number of genres mastered increased from 1 to 2, which demonstrates that compensations from listening skills began to decline and listening genres began to take over the role to keep the increase of test scores when two listening genres were mastered.

Table 5. ANOVA Multiple Comparisons for Increase of Genre Mastery.

| Increase of genre number | Average number of skills mastered | Test score means |
|--------------------------|----------------------------------|-----------------|
|                          | Difference | p    | Difference | p    |
| 0→1                      | 0.14176    | .354 | 0.80584    | .011 |
| 1→2                      | -0.41559   | .000 | 0.55687    | .024 |
| 2→3                      | 1.88732    | .000 | 5.53093    | .000 |

Both of the average number of listening skills mastered and the test score means increased drastically when the number of genres mastered increased from 2 to 3, which demonstrates that mastery of all listening genres has a strong contributory effect on mastering listening skills and thus pushes test scores even higher.

The researchers also compared the average numbers of the genres mastered when different numbers of listening skills were mastered, which is shown in Figure 3 accompanied with the corresponding average test scores.

The bars in Figure 3 represent the average numbers of listening genres mastered at different listening skill mastery situations. When no listening skill was mastered, the average number of listening genres mastered was 1.5 which was just at the moderate level. The average number of listening genres mastered increased and then decreased when the number of listening genres mastered increased from 0 to 3, but increased fast when the number of listening skills mastered increased all the way from 3 to 5. The subjects mastering only one listening skill even mastered more listening genres than those mastering four listening skills. The line in Figure 3 represents the change of test score means when the number of
listening skills mastered increased. When no listening skill was mastered, the test score mean was 11 which was not far below the moderate level. The test score means increased steadily when the number of listening skills mastered increased all the way from 0 to 5.

The researchers conducted One-Way ANOVA analyses to compare average numbers of listening genres mastered and test score means respectively at different listening skill mastery situations. Table 6 displays the multiple comparison results with the exclusion of the increase of skill number from 0 to 1 because only two subjects mastered zero skill.

The multiple comparison results of the One-Way ANOVA analyses were also based on the Games-Howell method because group sizes were quite different and equal variances were not assumed.

Table 6 also reflects that there may exist either compensatory or contributory relationships between listening skills and listening genre competence. The average number of listening genres mastered and the test score means were both around the moderate level when only one listening skill was mastered, which demonstrates that mastering a moderate number of listening genres compensates for the lack of listening skill mastery and plays an important role in maintaining test score means around the moderate level. When the number of skills mastered increased from 1 to 2, the average number of listening genres mastered decreased significantly. At the same time, the test score means increased significantly. The phenomenon demonstrates that compensations from listening genres began to decline and listening skills began to take over the role to keep the increase of test scores when two listening skills were mastered. When the number of skills mastered increased from 2 to 3, the average number of listening genres mastered decreased but not significantly. At the same time, the test score means increased significantly. The phenomenon demonstrates that the decline of listening genres came close to a stop and listening skills were ready to contribute to listening genre mastery when three listening skills were mastered.

Both of the average number of listening genres mastered and the test score means increased significantly when the number of skills mastered increased from 3 to 5, which demonstrates that mastering a majority of listening skills has a strong contributory effect on mastering listening genres and thus pushes test scores even higher.
According to the above elaborations on compensatory and contributory relationships, the researchers found that both listening skills and listening genre competence had compensatory and contributory effects on each other. Listening skills compensated in skill numbers when very few listening genres were mastered. Conversely, listening genre competence compensated in genre numbers when very few listening skills were mastered. Listening skills contributed to listening genre mastery when a majority of listening skills were mastered. Conversely, listening genre competence contributed to listening skill mastery when almost all listening genres were mastered. There was a process of decline of compensatory effects as more and more listening skills or listening genres were mastered. Whether compensatory or contributory, the relationship mechanism between listening skills and listening genre competence would always promote the increase of test scores.

Discussion

Genre Competence as Listening Attribute

By taking both listening skills and listening genre competences as cognitive attributes, cognitive diagnosis was conducted through the saturated G-DINA model. A Q-matrix containing both listening skills and listening genre competences was constructed through a validation process involving both statistical estimation and expert judgment. Based on that Q-matrix, good model fit was achieved for the CDM analysis, which demonstrates that listening skills and listening genre competences can be treated as cognitive attributes at the same time in the estimation of listening test data. The results of CDM research resonate with the theory of discoursal knowledge as one of the listening processes. Based on the empirical evidence, this study suggests that listening genre competence be taken into consideration in future research on listening skills.

Relationships Between Listening Skills and Genre Competence

Based on the distribution of the latent classes which is important output of CDM analysis, the researchers discovered prevailing coexistence relationships between listening skills and listening genre competence for those mastering the Lecture genre (010), the Dialog and Lecture genres (110), the Lecture and News genres (011), and all three genres (111), which demonstrates listening skills and listening genre competence, for the most part, are closely related. There still existed weak coexistence relationships for those mastering the Dialog genre (100) or the News genre (001) and conflictive relationships for those mastering the Dialog and News genres (101). Similar phenomenon was also discovered in the field of reading (Jang, 2009; Li et al., 2016). What is important is to avoid the weak coexistence or conflictive relationship in teaching and testing of listening comprehension.

How listening skills and listening genre competence are related involves an in-depth study on whether listening skills and listening genre competence are compensatory to each other. The obtained results of this study revealed that there also existed a contributory relationship between listening skills and listening genre competence, in addition to a compensatory relationship. Compensatory relationship existed when there was a lack of listening skill mastery or listening genre mastery. The subjects tended to master more listening skills than expected when very few listening genres were mastered. Conversely, they tended to master more listening genre competences than expected when very few listening skills were mastered. Contributory relationship existed when most of the listening skills or listening genres were mastered. Listening skills contributed to listening genre mastery when a majority of listening skills were mastered. Conversely, listening genre competence contributed to listening skill mastery when almost all listening genres were mastered.

Since compensatory relationship is usually related with low achievers and contributory relationship with high achievers, the transition from the compensatory relationship to the contributory relationship probably is related with a language threshold (Alderson, 1984) in the domain of listening. When learners’ listening proficiency does not pass the threshold, compensatory relationships prevail. When listening proficiency passes the threshold, contributory relationships prevail. By enhancing learners’ listening proficiency, the transition may come earlier.

Genre Competence and Listening Test Performance

By integrating test scores with genre mastery, the researchers found that among the subjects mastering only one genre, those who mastered Lecture genre got the highest average score (17.56 from 30). Those mastering a single Dialog genre got 11.1 and those mastering a single News genre got 11.0. Based on Table 4 which demonstrates the coexistence relationship between listening competence and listening genre competence, the researchers also found that 23.95% of the subjects mastered a single Lecture genre while those mastering a single Dialog genre or a single News genre merely accounted for 1.71% or less than 1% respectively.

The subjects mastering Lecture genre got higher scores probably because the attribute of Lecture genre coexists better with listening skill attributes. If the genre attribute cannot coexist with listening skill attributes well, the compensatory and contributory mechanism between genre competence and listening skills may not function well, which may lead to lower listening scores. In a word, whether mastery of a certain genre can lead to higher scores is probably determined by the coexistence relationship between listening skills and that genre.
Conclusion

Based on the latent class distribution which was important output of CDM analysis, the study analyzed three types of relationships between listening skills and genre competence. By analyzing how mastery of certain listening text genres goes with mastery of listening skills, the coexistence relationship was discovered. It was discovered that language learners tended to master more listening skills than expected so as to compensate for their lack of listening genres. Conversely, they tended to master more listening genre competences than expected so as to compensate for their lack of listening skills. It was also discovered that listening skills contributed to listening genre mastery when a majority of listening skills were mastered. Conversely, listening genre competence contributed to listening skill mastery when almost all listening genres were mastered.

All the three types of relationships between listening skills and genre competence may be conducive to learners’ listening performance. The phenomenon that the subjects mastering Lecture genre got higher scores probably can be explained by the fact that the attribute of Lecture genre coexists better with listening skill attributes. No matter whether compensatory or contributory relationship prevails, the listening scores invariably increase as shown in Figures 2 and 3. The transition between the two effects probably can be explained with the language threshold hypothesis which was first proposed in the research on reading (Alderson, 1984; Clarke, 1980). Learners’ listening proficiency or general L2 proficiency probably is the key to determining the point of transition from compensatory relationship to contributory relationship.

Teaching implications can also be derived from this study. Based on the empirical evidence, this study suggests that listening genre competence be taken into consideration in listening skill teaching and practice. If the listening genres involved are the three genres discussed in this research, importance should be attached to the teaching and practice of lecture genre because combinations of lecture genre with other genres have strong coexistence relationship with listening skills and the mastery of lecture genre, to a large extent, is a prerequisite for mastering dialog genre or/and news genre.

The research is also conducive to the design and compilation of listening tests. Test designers should try to adopt in listening tests different listening genres which have strong coexistence relationship with listening skills and avoid including in listening tests the listening genres which lead to weak coexistence and even conflictive relationship with listening skills so as to optimize the test performance. Since 2016, TEM4 examination has canceled the news section, which can also be seen as a move to optimize test performance by adopting appropriate listening genres.

Listening comprehension involves a complex cognitive process. Cognitive diagnosis can analyze the listening process at a finer-grained level of cognitive attributes and help to explore the skill-genre relationships. However, the study still has limitations. This study only investigated three listening text genres which are only a small portion of the varied listening genres. Further large sample research on the relationships between competence of other listening genres, such as movie and TV, and listening skills will be of more academic interest.

Authors’ Note

The data of this research were provided by the Committee of Test for English Majors in China. Jinsong Chen is also affiliated to University of Hong Kong, Hong Kong, China.

Acknowledgments

The Committee of Test for English Majors in China provided the researchers with test data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: National Social Science Foundation in China (17BYY101)

ORCID iD

Huilin Chen https://orcid.org/0000-0001-8040-3472

References

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem. In J. C. Alderson & A. H. Urquhart (Eds.), Reading in a foreign language (pp. 1–24). Longman.

Allison, D., & Tauroza, S. (1995). The effect of discourse organisation on lecture comprehension. English for Specific Purposes, 14, 157–173. https://doi.org/10.1016/0889-4906(95)00007-c

Bachman, L. F. (1990). Fundamental considerations in language testing. Oxford University Press.

Bazerman, C. (2009). Genre and cognitive development: Beyond writing to learn. In C. Bazerman, A. Bonini, & D. Figueiredo (Eds.), Genre in a changing world (pp. 279–294), Parlor Press and WAC Clearinghouse.

Berne, J. E. (1992). The effects of text type, assessment task, and target language experience on foreign language learners’ performance on listening comprehension (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

Bhatia, V. K. (1993). Analyzing genre: Language use in professional settings. Longman.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), Language and communication (pp. 2–27). Longman.

Alderson, J. C., & Widdowson, H. G. (1980). Reading in a foreign language: A reading problem or a language problem. In J. C. Alderson & A. H. Urquhart (Eds.), Reading in a foreign language (pp. 1–24). Longman.

Bhatia, V. K. (1993). Genre in a changing world. Parlor Press.

Clarke, J. E. (1992). Analyzing genre: Language use in professional settings. Longman.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), Language and communication (pp. 2–27). Longman.

Bhatia, V. K. (1993). Analyzing genre: Language use in professional settings. Longman.
Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1–47.

Carrell, P. L., & Eisterhold, J. C. (1983). Schema theory and ESL pedagogy. *TESOL Quarterly, 17*, 553–573. https://doi.org/10.2307/3586613

Chen, H., & Chen, J. (2016). Exploring reading comprehension skill relationships through the G-DINA model. *Educational Psychologist, 36*, 1049–1064. https://doi.org/10.1080/00144330.2015.1076764

Chen, J. (2017). A residual-based approach to validate Q-matrix specifications. *Applied Psychological Measurement, 41*, 277–293. https://doi.org/10.1177/0146621616686021

Chen, J., de la Torre, J., & Zhang, Z. (2013). Relative and absolute fit evaluation in cognitive diagnosis modeling. *Journal of Educational Measurement, 50*, 123–140. https://doi.org/10.1111/j.1745-3984.2012.00185.x

Clarke, M. A. (1980). The short circuit hypothesis of ESL reading-or when language competence interferes with reading performance. *Modern Language Journal, 64*(2), 203–209. https://doi.org/10.1111/j.1540-4781.1980.tb05186.x

Connor, U., & Mbaye, A. (2002). Discourse approaches to writing assessment. *Annual Review of Applied Linguistics, 22*, 263–278. https://doi.org/10.1017/s0267190502000144

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika, 76*, 179–199. https://doi.org/10.1007/s11336-011-9207-7

English Major Division of National Foreign Languages Advisory Board. (2000). *English teaching syllabus for English majors [Chinese]*. Foreign Language Teaching and Research Press.

Field, J. (2013). Cognitive validity. In A. Geranpayeh & L. Taylor (Eds.), *Examining listening: Research and practice in assessing second language listening* (pp. 77–151). Cambridge University Press.

Flowerdew, L. (2000). Using a genre-based framework to teach organizational structure in academic writing. *ELT Journal, 54*(4), 369–378. https://doi.org/10.1093/elt/54.4.369

Francis, H., & Hallam, S. (2000). Genre effects on higher education students’ text reading for understanding. *Higher Education, 39*, 279–296.

Freedman, A., & Medway, P. (Eds.). (1994). *Learning and teaching genre*. Greenwood.

Gierl, M. J., Leighton, J. P., & Hunka, S. M. (2005). An NCME instructional module on exploring the logic of Tatsuoka’s rule-space model for test development and analysis. *Educational Measurement Issues and Practice, 19*(3), 34–44. https://doi.org/10.1111/j.1745-3992.2000.tb00036.x

Goh, C. C. M., & Aryadoust, V. (2015). Examining the notion of listening subskill divisibility and its implications for second language listening. *International Journal of Listening, 29*(3), 109–133. https://doi.org/10.1080/10904018.2014.936119

Henry, A., & Roseberry, R. L. (1998). An evaluation of a genre-based approach to the teaching of EAP/ESP writing. *TESOL Quarterly, 32*, 147–156.

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly, 30*(4), 693–722. https://doi.org/10.2307/3587930

Jang, E. E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for fusion model application to LanguEdge assessment. *Language Testing, 26*, 031–073. https://doi.org/10.1177/0265532208097336

Johnson, K., & Johnson, H. (Eds.). (1999). *Encyclopedic dictionary of applied linguistics: A handbook for language teaching*. Blackwell Publishers. https://doi.org/10.1111/b.9780631214823.1999.x

Li, H., Hunter, C. V., & Lei, P. W. (2016). The selection of cognitive diagnostic models for a reading comprehension test. *Language Testing, 33*(3), 391–409. https://doi.org/10.1177/0265532215590848

Littlefair, A. (1991). *Reading all types of writing: The importance of genre and register for reading development*. Open University Press.

Liu, J., & Wu, S. (2019). Research on inventory of China standards of English [Chinese]. Higher Education Press.

Manzouri, H. A. (2015). The effects of genre-based instruction on Iranian EFL learners’ listening comprehension. *Research in English Language Pedagogy, 3*(1), 74–82.

Martin, J. R. (1985). *Factual writing: Exploring and challenging social reality*. Deakin University Press.

McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a second language. *Language Learning, 36*(2), 109–123. https://doi.org/10.1111/j.1467-1770.1986.tb00374.x

Metruk, R. (2019). Using English movies and TV programs for developing listening skills of EFL learners. *Information Technologies and Learning Tools, 70*(2), 227–236. https://doi.org/10.33407/itlt.v70i2.2488

Miller, C. R. (1984). *Genre as social action*. *Quarterly Journal of Speech, 70*(2), 151–167.

Min, S., & Xiong, L. (2019). Exploring the compensatory mechanism of EFL listening ability: A cognitive diagnostic assessment approach. *Linguistics and Applied Linguistics [Chinese], 02*, 254–266.

Myskow, G., & Gordon, K. (2010). A focus on purpose: Using a genre approach in an EFL writing class. *ELT Journal, 64*(3), 283–292. https://doi.org/10.1093/elt/ccp057

North, B. (2007). The common European framework of reference: Development, theoretical and practical issues. *babylonia, 1*, 22–29.

Nunan, D. (2008). Exploring genre and register in contemporary English. *English Today, 24*(2), 56–61. https://doi.org/10.1017/s0266078408000217

Nushi, M., & Orouji, F. (2020). Investigating EFL teachers’ views on listening difficulties among their learners: The case of Iranian context. *Sage Open, 10*(2), 1–16. https://doi.org/10.1177/2158244020917393

Peterson, P. W. (1991). A synthesis of methods for interactive listening. *Teaching English as a Second or Foreign Language, 2*, 106–122.

Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*(2), 219–240.

Sawaki, Y., Kim, H. J., & Gentile, C. (2009). Q-Matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessments. *Language Assessment Quarterly, 6*(3), 190–209. https://doi.org/10.1080/15434300902801917

Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing, 8*(1), 23–40. https://doi.org/10.1177/026553229100800103
Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Thompson, S. (1994). Frameworks and contexts: A genre-based approach to analysing lecture introductions. *English for Specific Purposes, 13*, 171–186. https://doi.org/10.1016/0889-4906(94)90014-0

Toledo, P. F. (2005). Genre analysis and reading of English as a foreign language: Genre schemata beyond text typologies. *Journal of Pragmatics, 37*(7), 1059–1079. https://doi.org/10.1016/j.pragma.2005.01.002

Valette, R. (1977). *Modern language testing*. Harcourt Brace Jovanovich.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*, 191–210. https://doi.org/10.1017/s0261444807004338

Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.

Yang, M., & Liu, J. (2019). China’s standards of English language ability and business English testing and assessment [Chinese]. *Foreign Languages in China, 16*(3), 13–20.

Yi, Y. S. (2017). Probing the relative importance of different attributes in L2 reading and listening comprehension items: An application of cognitive diagnostic models. *Language Testing, 34*(3), 337–355. https://doi.org/10.1177/0265532216646141

Young, L. (1995). University lectures – Macro-structure and micro-features. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 159–176). Cambridge University Press.