HOW
Volume 31, Number 1, pages 37 - 66
https://doi.org/10.19183/how.31.1.694

# Exploring Listening Assessment in a Colombian EFL Context: A Case Study

## Explorar la Evaluación de la Comprensión de Escucha en un Contexto EFL Colombiano: Un Estudio de Caso

**Ana Gutiérrez-Rojas**[1]
**Nayibe Rosado-Mendinueta**[2]
Universidad del Norte, Barranquilla, Colombia.

### Abstract

Listening comprehension is a crucial skill in English as a Foreign Language education, yet listening assessment still needs to be explored. This study aims to understand how listening is assessed in a Colombian private language institution and its potential connection to students' underperformance in listening proficiency tests. We characterize the listening assessment used in B1-level classes through a descriptive case study. Utilizing a rubric, we analyzed fourteen listening tasks from various sources, including course materials and a mock PET exam. Our analysis had two primary objectives: (a) to uncover the rationale behind test design and its alignment with the curriculum, and (b) to evaluate the knowledge types targeted and the forms of listening assessment employed using categorization. Our findings reveal that despite the curriculum's holistic listening development goals, assessments predominantly focus on phonological knowledge through dictation tests, in which students primarily engage with audio media and recordings. These findings suggest a misalignment in the way listening is approached during the whole course, the examinations used to assess listening during the B1 level, and

[1] She is an English language teacher and a master's student at Universidad del Norte in Colombia. Her research interests include L2 listening comprehension and assessment.
rojasana@uninorte.edu.co
ORCID: https://orcid.org/0000-0003-3950-4487

[2] She is a language teacher and teacher educator at Universidad del Norte in Colombia. She is a co-founder and immediate past president of the Latin American Association of Language Testing and Assessment( LAALTA).
nrosado@uninorte.edu.co
ORCID: https://orcid.org/0000-0003-1865-2464

Received: October 1st, 2021. Accepted: February 5th, 2023.

the use of an external standardized test as the exit exam. This study holds potential implications for curriculum alignment and the enhancement of language assessment literacy within our context, shedding light on potential factors contributing to students' underperformance in listening comprehension.

*Keywords*: listening assessment, listening activities, curricular alignment, language education, EFL context, language proficiency

**Resumen**

La comprensión de escucha es una destreza crucial en la enseñanza del inglés como lengua extranjera, sin embargo, la evaluación de la comprensión auditiva aún necesita ser explorada. El objetivo de este estudio es entender cómo se evalúa la comprensión auditiva en una institución privada de idiomas colombiana y su posible relación con el bajo rendimiento de los estudiantes en las pruebas de competencia auditiva. Caracterizamos la evaluación de la comprensión auditiva utilizada en las clases de nivel B1 a través de un estudio de caso descriptivo. Utilizando una rúbrica, analizamos catorce tareas de comprensión auditiva de diversas fuentes, incluidos los materiales del curso y un simulacro de examen PET. Nuestro análisis tenía dos objetivos principales: a) descubrir la justificación del diseño de las pruebas y su alineación con el plan de estudios, y b) evaluar los tipos de conocimiento a los que se dirigen y las formas de evaluación de la comprensión auditiva empleadas utilizando una categorización. Nuestros hallazgos revelan que, a pesar de los objetivos holísticos de desarrollo de la escucha del plan de estudios, las evaluaciones se centran predominantemente en el conocimiento fonológico a través de pruebas de dictado, donde los estudiantes se involucran principalmente con medios de audio y grabaciones. Estos hallazgos sugieren un desajuste en la forma en que se aborda la comprensión oral durante todo el curso, los exámenes utilizados para evaluar la comprensión oral durante el nivel B1 y el uso de una prueba estandarizada externa como examen de salida. Este estudio informa sobre la alineación del currículo y el desarrollo de la competencia en evaluación del lenguaje en nuestro contexto, arrojando luz sobre posibles razones para el bajo rendimiento de los estudiantes en la comprensión auditiva.

*Palabras clave*: evaluación de escucha, actividades de escucha, alineación curricular, educación en lengua, inglés como lengua extranjera, suficiencia en lengua extranjera

## Introduction

Listening comprehension is a cornerstone of second language acquisition, particularly in English as a Foreign Language (EFL) education. However, despite this skill's significance, the listening assessment remains an area with substantial room for exploration and improvement. In the Colombian EFL context, listening assessment practices warrant closer examination, especially considering the persistent underperformance of students in listening proficiency tests.

This study is driven by the need to bridge the gap between the curriculum's intentions, which emphasize a holistic approach to listening development, and the actual outcomes observed in students' performance when facing proficiency tests, such as the B1 Preliminary English Test (PET). Institutional data have shown a notable inconsistency between students'

expected and actual listening proficiency levels. Despite the curriculum's strong emphasis on communicative competence, students often fall short of the expected standards, as reflected in their scores on the Cambridge English Qualifications mock exams.

This study aligns with existing literature that underscores the importance of understanding how listening is assessed in specific educational contexts (Rost, 2011). As mentioned earlier, the study responds to the persistent challenge of students' underperformance in listening proficiency tests, a concern that scholars have addressed in various settings.

Morales and Beltrán's (2006) action research study emphasized the significance of materials selection and systematic lesson plans for improving students' listening comprehension. Similarly, Córdoba-Zúñiga and Rangel-Gutiérrez (2018) emphasized meaningful oral tasks and Task-Based Language Teaching to promote listening fluency and highlight the value of engaging activities. Mayora's (2017) exploration of extensive listening highlighted the potential for self-assessment and self-regulation, reflecting the importance of metacognitive strategies to improve students' listening skills. Moreover, Ballesteros-Muñoz and Tutistar (2014) pointed to the positive impact of teaching specific, measurable, achievable, relevant, and time-bound (SMART) goal setting on listening comprehension and self-efficacy, reinforcing that assessment methods can influence learner performance and attitudes.

Our focus on understanding the rationale behind test design and its alignment with the curriculum resonates with the literature's emphasis on the need for materials beyond literal comprehension (Cárcamo-Morales, 2018). The analysis of listening tasks' knowledge types and forms of assessment using Rost's (2011) categorization aligns with the broader interest in exploring various dimensions of listening assessment (Hernández-Ocampo & Vargas, 2013; Sevilla-Morales & Chaves-Fernandez, 2019), including teachers' language assessment literacy, which refers to the knowledge, skills, and principles needed for contextualized language assessment (Giraldo et al., 2023). This is an emerging theme in the Colombian EFL field (Giraldo, 2018, 2021a, 2021b; Giraldo & Murcia, 2018, 2019).

This study contributes to the growing body of research on listening assessment practices in EFL contexts, shedding light on the possible reasons behind students' underperformance. Our focus is to delve into the characteristics of listening assessment within a specific educational context: a private language institution in Colombia. This institution aligns its programs with the Common European Framework of Reference (CEFR, Council of Europe, 2001) and offers CEFR-aligned courses, including the B1 level course where this study was embedded. The course consisted of two modules, each comprising six units based on the Navigate Pre-intermediate B1 coursebook. The assessment framework includes periodic unit-level assessments and a final exit exam, the Preliminary English Test (PET) mock exam.

Concerning students' performance in listening proficiency tests, the study poses the following research question: "What are the characteristics of the listening assessment at the B1-level EFL course at a private language institute in Colombia?"

The subsequent sections of this article provide a framework for understanding the nature of listening and introduce key assessment concepts. The methodology section details the research design, data collection, and analysis procedures to investigate the identified gap. Finally, the findings and discussion section illustrates the characteristics of listening assessment in this EFL context, with potential implications for curriculum alignment and the development of teachers' assessment literacy.

## Conceptual References

The conceptual reference section consists of three key elements: Listening, Language Assessment Principles, and Listening Assessment. By incorporating these elements, we hope to provide a clear path for exploring listening assessment practices and their alignment with the complexities of listening in the study context.

### *Listening*

Listening is often related to the faculty of the ears to perceive sounds; therefore, the study of the physiological components has been approached by audiologists (Worthington & Bodie, 2017). In contrast, communication scholars have studied the individual and relational components of listening. For these researchers, listening is "a multidimensional construct that consists of complex (a) affective processes, such as being motivated to attend to others; (b) behavioral processes, such as responding with verbal and nonverbal feedback; and (c) cognitive processes, such as attending to, understanding, receiving, and interpreting content and relational messages" (Worthington & Bodie, 2017, p. 3). This concept of listening includes an extensive set of processes. In a similar line, Rost (2011) provides an ample definition whereby "listening is an integrated ability that requires several overlapping psycholinguistic abilities" (p. 117). These definitions point to listening subprocesses (Rost, 2011; Worthington & Bodie, 2017), as represented in Figure 1.



Figure 1 portrays listening as a complex process that encompasses four other processes. So, in the present study, the definition of listening entails that listening is a dynamic and complex process that encompasses four other processes - neurological processing, linguistic processing, semantic processing, and pragmatic processing - in which meaning is co-constructed from the perception of the sounds of a language.

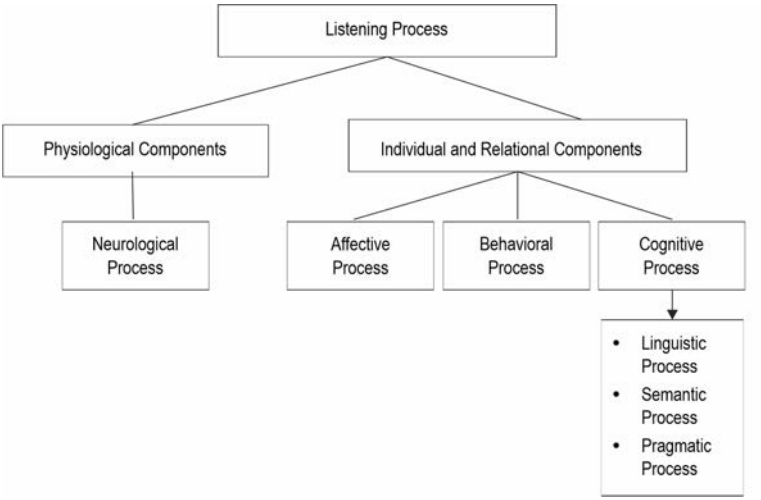


**Figure 1**. Listening and its Processes

**Note**. The main components and processes of the listening process. It is adapted from Rost (2011) and Worthington and Bodie (2017).

## *Language Assessment Principles*

In language assessment, foundational principles guide the creation and application of assessment instruments to ensure their effectiveness. These principles include validity, practicality, reliability, authenticity, and washback (Brown, 2004; Brown & Abeywickrama, 2010; Malone, 2011). Similarly, Bachman and Palmer (1996) identify six properties for quality control in test development: reliability, construct validity, authenticity, interactiveness, impact, and practicality, with usefulness as their primary function.

Validity, often deemed the most critical principle, ensures the accurate measurement of the intended construct through content-related, criterion-related, construct, and face validity (Brown & Abeywickrama, 2010). Construct validity, as emphasized by Buck (2001), is often regarded as the most pivotal property, as it ensures that a test accurately measures the intended construct, serving as the foundation for test usefulness. Practicality addresses logistical considerations, encompassing cost, time, ease of administration, scoring, and interpretation (Brown, 2004; Mousavi, 2009). Reliability hinges on the consistency of test results across various factors, requiring clear instructions, uniform rubrics, and unambiguous items or tasks (Bachman & Palmer, 1996; Brown & Abeywickrama, 2010). Authenticity stresses the importance of assessment tasks mirroring real-world language use (Bachman & Palmer, 1996; Brown & Abeywickrama, 2010). The washback effect underscores the influence of

assessments on teaching, learning, motivation, materials, and classroom practices (Hughes, 2003; McKinley & Thompson, 2018; Rost, 2011; Alderson & Wall, 1993).

Brown and Abeywickrama (2010) and Bachman and Palmer (1996) provide comprehensive language assessment frameworks, sharing fundamental principles while differing in some definitions and perspectives. Both prioritize reliability and practicality but vary in their interpretation of validity. Brown and Abeywickrama emphasize various types of validity evidence, while Bachman and Palmer focus on construct validity as the foremost quality. Additionally, they diverge on the definitions of impact and washback, with Brown and Abeywickrama emphasizing the effects of tests on teaching and learning and Bachman and Palmer encompassing broader social consequences, including policy and societal values. Moreover, Brown and Abeywickrama introduce face validity as a principle, emphasizing its role in gaining acceptance and confidence among stakeholders, whereas Bachman and Palmer do not consider it a valid criterion, focusing on test effectiveness (Brown & Abeywickrama, 2010; Bachman & Palmer, 1996).

### *Listening Assessment*

Defining the construct of listening is a fundamental aspect of assessment, as it ensures that the assessment aligns with the intended goals and accurately measures the targeted skills. Green (2019) underscores the need to base the construct on various sources, including the school curriculum, national standards, and CEFR descriptors. This alignment is essential for collecting evidence that validates the definition of the listening construct.

Buck (2001) further emphasizes the central role of construct validity in assessment. He distinguishes between two types of knowledge crucial for language comprehension: linguistic knowledge, which encompasses sounds, words, grammar, semantics, and more; and non-linguistic knowledge, which relates to aspects like context, topic, and cultural elements. Depending on their application in different situations, these knowledge types underpin the bottom-up and top-down listening views.

Rost (2011) contributes to this discourse by presenting a comprehensive map of listening ability, highlighting its overlap with general language ability. This map incorporates five key knowledge domains: general, pragmatic, syntactic, lexical, and phonological. It underscores how these knowledge areas intersect, further emphasizing their importance in the context of listening assessment. Figure 2 below depicts such knowledge.



Additionally, Rost (2011) offers practical guidance for discussions on listening assessment. He categorizes knowledge types (See Appendix A and B and Figure 2 above) and forms of listening assessment, which are the different ways of measuring and evaluating

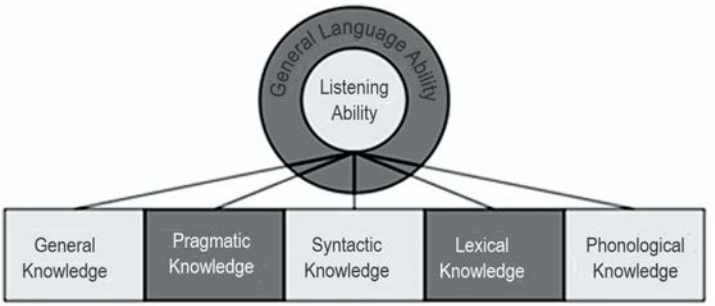


**Figure 2**. General Language Ability and Listening Ability

**Note**. From Teaching and Researching Listening (p. 212) by Rost (2011).

the listening ability of language learners. He proposes four primary forms of listening assessment: discrete-point, integrative, communicative, and performance-based.

Discrete-point listening assessment dissects listening into isolated components like phonology, lexis, syntax, and semantics, employing question types such as multiple-choice and true-false. While straightforward to administer and score, it lacks the realism of real-life listening contexts. Secondly, integrative listening assessment acknowledges the interplay between various components and evaluates them about one another, employing methods like cloze exercises and dictation. This approach offers a more authentic assessment but remains constrained by test format and content. Thirdly, communicative listening assessment focuses on the dynamic exchange of meaning and intention in listening, testing the ability to employ listening for diverse purposes in varied contexts through tasks such as role-plays and problem-solving activities. It offers greater realism and interaction but can be challenging to design and score. Finally, performance-based listening assessment emphasizes listening to achieve specific goals or outcomes, evaluating learners' capacity to apply listening in actual or simulated settings using tasks like presentations or projects. This highly valid and meaningful approach demands more complexity and subjectivity in assessment design (see Appendix C and D). These resources serve as valuable tools to design assessments that align with the complexities of listening.



The course syllabus and assessment package are based on the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001); we also used it to analyze the communicative listening activities. They are defined as those involving the reception and processing of oral or audio-visual input for various purposes and functions in different contexts and situations. These activities are categorized into four types: listening as a member of a live audience, listening to media and recordings, listening as a learner, and

listening for professional purposes. Listening as a live audience member entails engaging with live presentations and reacting appropriately to the speaker's message. Listening to media and recordings involves understanding recorded or broadcast input. Listening as a learner focuses on acquiring knowledge and skills from learning materials. Listening for professional purposes involves listening related to one's work or occupation, often requiring specific tasks and goals.

Based on these foundational concepts, we proceed to analyze the tests used to assess students' listening skills, as described in the following section.

## Methodology

In this section, we outline the methodology employed in our study, beginning with an introduction to the research design and its contextual background. We then describe the data collection and analysis procedures to meet the study's objectives.

A descriptive study is a type of research that aims to describe the characteristics of a population or phenomenon. It can answer questions such as what, where, when, and how, but not why. A descriptive study can use various methods, such as surveys, observations, and case studies to collect data that can be analyzed for frequencies, averages, patterns, and categories. A descriptive study is helpful when not much is known about the phenomenon to provide a clear and detailed picture of said phenomenon (Islam & Aldaihani, 2022). A case study "allows the researcher to examine a case in depth in its real context" (Stake, 2005; Yin, 2013). An example can be a person, a community, an event, or a policy for further understanding (Bhattacharya, 2009; Merriam, 1988; Stake, 2005).

The present study utilizes a descriptive case study design to characterize the listening assessment practices employed by EFL teachers in B1-level classes at a private language institution in the Northeastern region of Colombia. The institution offers a CEFR-aligned language program, encompassing A1, A2, B1, and B2. Each course is designed to equip learners with reading, writing, listening, and speaking skills, facilitating their progression to the targeted proficiency level. Most students enrolled in these courses are young adults, and the teaching staff holds bachelor's degrees in English education. Several teachers have further validated their teaching competence by completing the Knowledge of Teaching (TKT) certification. Additionally, all instructors possess language proficiency certificates, validating their English proficiency levels at or above B2.



### *Study Context*

The study unfolds within the framework of a B1-level course, structured into two modules, each encompassing six units sourced from the course book "Navigate Pre-

Intermediate B1." Over the course duration, learners are evaluated after each unit, with an additional checkpoint assessment following Unit 6. This checkpoint entails a midterm examination from the curriculum testing package (CCTP). Upon completing all twelve units in the B1-level coursebook, students undergo a mock examination derived from the Cambridge English Main Suite, specifically, the Preliminary English Test (PET), serving as a final exit exam within the context of an external testing package (ETP).

### *Study Objectives*

The primary objective of this research is to characterize the listening assessment methods employed in B1-level classes at the private language institution. This characterization aims to illuminate potential reasons underlying students' consistent underperformance in listening proficiency assessments, notably the PET. Consequently, our study analyzes listening examinations utilized in the B1-level course.

### *Data Collection and Analysis*

The characterization process consists of two fundamental steps. Initially, we gathered established examinations pertinent to the B1 course, with a meticulous review conducted to ascertain the inclusion of listening sections and to quantify the number of tasks within each examination. Subsequently, guided by Buck's (2001) specifications for test validators, we analyzed each type of examination to elucidate the rationale behind its design and its alignment with the instructional curriculum.

The second step involved a granular analysis of individual listening tasks, primarily driven by Rost's (2011) categorization of listening abilities. Given that the program syllabus and course content are rooted in the CEFR framework, we employed the CEFR's descriptions of communicative listening activities to identify the specific types of listening activities integrated into each task. This multifaceted analysis unfolded across three dimensions:

- Characterizing the type of knowledge assessed by each instrument.
- Identifying the forms of listening assessment embedded within each task.
- Categorizing the type of communicative listening activities entailed in each instrument.



### *Analysis of Listening Tasks*

To facilitate our analysis, we designed a checklist aligned with Rost's (2011) categorization of general language abilities and listening-specific attributes. We initially applied this checklist to assess the curriculum-based assessments and subsequently to the four subsections of the mock PET examination (refer to Appendix A and B).

Furthermore, we scrutinized the forms of listening assessment embedded in each task, guided by Rost's (2011) description of assessment formats. Details of this process are presented in Appendix C and D, which encompass the checklists created to identify the various forms of listening assessment evident in the tests and examinations utilized in the B1-level course. To complete our analysis, we meticulously reviewed each task to determine the associated communicative listening activities, referring to the CEFR's classifications.

### *Ethical Considerations*

Before initiating the study, we upheld rigorous ethical considerations. Permission to access the requisite data was diligently sought from the institution's director, accompanied by a comprehensive explanation of the study's purpose and the potential benefits it could offer the institution. Of notable importance is the involvement of Researcher 1, who maintains a direct affiliation with the institution and plays a pivotal role in facilitating communication and collaboration. All participating teachers and relevant stakeholders were informed about the research objectives, and their written consent was duly obtained. The director, recognizing the persistent challenges faced by students in listening comprehension, expressed keen interest in the study. Throughout the research process, ethical principles and practices were meticulously adhered to, safeguarding the privacy and rights of all participants.

## Findings and Discussion

This section presents the findings concerning the specifications of examinations utilized in the B1-level course at a private language institution. We comprehensively analyzed two types of examinations: (a) the unit and final tests derived from the course curriculum testing package (CCTP) and (b) the mock PET exam selected as the B1-level exit assessment. We summarize the findings and engage in a thorough discussion to shed light on the implications for students' listening skills and test performance.



### *Summary of Findings*

Table 1 offers a comprehensive overview of the specifications of the examinations at the B1 level, categorizing them into CCTP and ETP (PET). It outlines the purpose, theoretical framework, listening construct, and alignment with test goals for each examination.

**Table 1.** Specifications of the Examinations at the B1 level

| Specification | CCTP | | ETP |
|---|---|---|---|
| | **Tests** | **Midterm Exam** | **PET Exam** |
| **Purpose of the test** | Unit Test | Proficiency test | Proficiency test |
| **Description of Theoretical Framework** | Each unit test measures the understanding of the listening decoding skills presented in the unit in a similar context to the one. | Aligning language to the CEFR. They resembled established international English language exams not only in their form but also in their purpose. | Students should be able to understand and respond to public announcements, show precise understanding of short factual utterances to make identifications based on these, and extract information from speech that will contain redundancies and language outside the defined limits of PET. |
| **Listening construct** | Relying on Field's research, listening in Navigate focuses on features of the spoken language. Therefore, it focuses on word recognition and decoding skills for listening to later switch attention from it to building up the speaker's purpose and the conversation. | | It is a complex skill operating at several levels and must be practiced accordingly (Field, 2008). |
| **Why does it meet the purpose of the test?** | Only listening sections to test the listening decoding skills covered in the class unit. | Listening tasks look like tasks in PET. | Assessment of candidates' ability to understand dialogues and monologues in both informal and neutral settings on a range of everyday topics. |

**Note**. Information about CCTP was taken from Navigate: Pre-intermediate B1: Teacher's Guide with Teacher's Support and Resource Disc. (p. 22-23, 32-33) by Merifield et al. (2015). Oxford University Press. Information about ETP, the Preliminary English Test (PET), was taken from the Preliminary English Test: Handbook. (p. 35) (2005). University of Cambridge ESOL Examinations.

The analysis reveals distinctions between the examinations. The unit tests within the CCTP primarily gauge students' comprehension of listening decoding skills relevant to the specific unit context. In contrast, the midterm and PET exams aim to assess learners' proficiency at the B1 level.

The discrepancies in the goals and constructs of these examinations prompt a significant discussion. As an integral component of the CCTP, the unit tests emphasize word recognition and decoding skills, with a gradual shift towards a more comprehensive understanding of the

speaker's intent and overall conversation context. This approach aligns with Field's signal-based approach (Field, 2003), emphasizing progression from perception to comprehension.

Conversely, as outlined in the Preliminary English Test: Handbook (2005), the PET exam focuses on evaluating students' ability to understand dialogues and monologues in diverse informal settings. The disparity in examination constructs becomes evident when comparing the number of listening tasks: the PET exam consists of four listening tasks, whereas the midterm exam includes only two.

Including the PET exam as the exit assessment aims to validate that the program attains the expected English proficiency level of a B1 student according to the CEFR. Nevertheless, the course curriculum and its associated examinations emphasize the development of word recognition and decoding skills, assuming that this training will ultimately enhance learners' listening proficiency at the anticipated level and facilitate success in the PET exam.

However, the analysis echoes the findings of Jiménez et al., (2017), who stress the importance of explicit test preparation and consistent self-evaluation processes in bridging the gap between standardized test results and curriculum objectives. While the midterm exam aligns with standardized test tasks to some extent, its limited number of listening tasks and the absence of comprehensive test preparation may explain students' lower-than-expected scores in the PET exam's listening section.

Furthermore, implementing consistent and continuous self-evaluation processes, as Jiménez et al., (2017) propose, can facilitate ongoing analysis of the relationship between standardized test results and curriculum objectives. This approach will empower teachers to make informed decisions to enhance students' language skills and performance on standardized tests.

### *CCTP and ETP Assess Different Types of Knowledge Regarding Listening*

The analysis uncovers a notable diversity in the types of knowledge assessed by examinations from the CCTP and ETP (PET). The curriculum test packaging at the B1 level primarily addresses phonological knowledge related to understanding English language sounds in fast speech and recognizing spoken words and lexical phrases (see Appendix A). This alignment is consistent with the identified purpose of the unit tests from the CCTP, focusing on the evaluation of listening decoding skills (see Table 2 above).



However, Cárcamo-Morales (2018) highlights the importance of fostering active interaction with audio texts, framing listening as a process that involves decoding information and organizing, evaluating, and responding to it. He asserts that limiting students' opportunities for active engagement with aural texts and neglecting cognitive processes

associated with comprehension could hinder their performance in standardized tests. This observation is particularly relevant to the curriculum's exclusive emphasis on phonological knowledge.

Within the curriculum, only one task in test number six assesses syntactic knowledge, explicitly targeting the understanding of cohesion markers in discourse (see Figure 3 below). This limited inclusion of syntactic knowledge assessment suggests room for enhancement in aligning assessments with a more comprehensive understanding of listening skills.

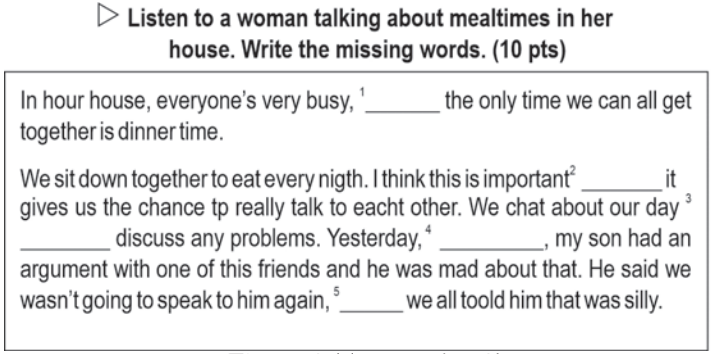


**Figure 3**. Test Number Six

Compared to the tests following each unit, the listening tasks in the midterm exam extend their focus beyond the recognition of spoken words and lexical phrases. Both tasks in the exam evaluate pragmatic knowledge, particularly the ability to follow the flow of given information versus new information. Our identification of the test specifications (Table 2) suggests that the midterm exam's tasks resemble established international English language exams (Merifield et al., 2015) and aim to assess similar knowledge categories as outlined in Appendix A.

Analyzing each section of the PET listening exam reveals a distinct approach. While the curriculum tests and midterm exams primarily emphasize phonological knowledge to test listening decoding skills, the PET exam integrates multiple dimensions of language knowledge, including phonological, lexical, pragmatic, and general knowledge (see Appendix B). This comprehensive approach aligns with the PET exam's purpose of assessing the learner's ability to perform tasks corresponding to the CEFR's description of the B1 level (Appendix E).



The divergence in the types of knowledge assessed by examinations from the CCTP and ETP raises significant considerations. Latimer (2009) observed a similar challenge in

his study on the washback effects of PET at an Argentinean bilingual school. Although the school adopted the PET as an external, credible means of demonstrating students' English proficiency, teachers expressed reservations, believing that the PET needed to encompass the full scope of students' English knowledge and abilities.

Considering these findings, Latimer (2009) suggested integrating PET preparation into extracurricular activities to avoid disrupting the institution's holistic communicative language mission. Similarly, the analysis of our assessments implies that examinations from the CCTP and ETP evaluate different dimensions of listening knowledge. Consequently, the institution should consider aligning its curriculum objectives, examination constructs, and test preparation strategies to address these disparities comprehensively.

### *Lack of Variety in Listening Assessment Forms*

The analysis has revealed a significant limitation in the forms of listening assessments used within the B1-level course. Notably, all unit tests feature the same integrative test format, akin to a complete or partial dictation, primarily focusing on scoring based on the correct supply of missing words (Rost, 2011, p. 216). This lack of variety in assessment forms raises concerns about the scope of listening assessment in unit tests.

Cárcamo-Morales (2018) articulated a similar concern when analyzing tasks in an English text. While different types of tasks were present, they predominantly practiced the same type of knowledge. This pattern restricts students from progressing to more intricate levels of listening comprehension. In the context of the analyzed listening tests, the lack of variety in form and the type of knowledge assessed limits the breadth of listening assessment throughout the unit tests. Furthermore, it hinders learners' exposure to diverse task formats commonly encountered in international exams, such as the PET.

### *Lack of Exposure to Different Types of Communicative Listening Activities in CCTP*

The analysis of the course's curriculum and associated assessments revealed a notable limitation—students' limited exposure to diverse communicative listening activities. We assessed the types of communicative "listening activities" (Council of Europe, 2001, p. 65) presented in each test from the curriculum. As depicted in Table 3 below, listening to audio media and recordings appears consistently across tests, even within the midterm exam. Furthermore, some activities, such as sentences with gaps to be filled, were identified in tests with two listening activities. Nevertheless, these activities do not correspond to any of the communicative listening activities described by the CEFR.

**Table 3**. A Checklist for the Communicative Listening Activities Addressed in CCTP

| Communicative Listening Activities | Test 3 Pt 1 | Test 3 Pt 2 | Test 4 | Test 6 | Test 8 Pt 1 | Test 8 Pt 2 | Test 11 Pt 1 | Test 11 Pt 2 | Exam Part 1 | Exam Part 2 |
|---|---|---|---|---|---|---|---|---|---|---|
| Listening to public announcements and instructions | — | | | | — | | — | | | |
| | | | | | | | | | | |
| Listening to audio media and recordings | — | ✓ | ✓ | ✓ | — | ✓ | — | ✓ | ✓ | ✓ |
| Listening as a member of a live audience | — | | | | — | | — | | | |
| Listening to conversations between native speakers | — | | | | — | | — | | | |

These findings parallel those of Morales and Beltran (2006), who observed that most listening materials in coursebooks engage learners in artificial, nearly flawless language use—an approach considered beneficial by designers. However, the authors highlight the importance of exposing learners to authentic listening situations reflecting real-life language use. Their study found that natural speech activities, such as films, news, and cartoons, posed more significant difficulties for students. This underscores the need for diversified listening experiences.

In contrast, our PET-exam analysis revealed various communicative listening activities integrated into its sections (Table 4). This diversity likely stems from the PET's alignment with the principles and approach of the CEFR, as it aims to assess English proficiency.

**Table 4**. A Checklist for the Communicative Listening Activities Addressed in ETP

| Communicative Listening Activities | Part 1 | Part 2 | Part 3 | Part 4 |
|---|---|---|---|---|
| Listening to public announcements and instructions | ✓ | | ✓ | |
| | | | | |
| Listening to audio media and recordings | ✓ | ✓ | | |

| Communicative Listening Activities | Part 1 | Part 2 | Part 3 | Part 4 |
|---|---|---|---|---|
| Listening as a member of a live audience | | | | |
| Listening to conversations between native speakers | ✓ | | | ✓ |

This disparity in communicative listening activities between the course curriculum and the PET exam has several implications. First, while students take the PET exam upon completing the course curriculum for the B1 level, an achievement test tailored to assess the skills taught within the curriculum might be more appropriate. Such a test could provide insights into the effectiveness of the listening approach outlined in the curriculum.

On the other hand, if the PET exam is intended to serve as an external measure of students' English proficiency, like Latimer's (2009) findings, a different approach is needed. PET preparation should be integrated into the curriculum, making it a fundamental component of course examinations.

Our analysis has highlighted the disconnect between the course curriculum's approach to listening, which emphasizes word recognition until it becomes automated, and the expectation that students demonstrate B1-level listening abilities in an external examination, such as the PET. Interestingly, an examination of students' results consistently reveals the lowest scores in the listening section (see Appendix F).

We can say that assessment practices uncovered in this study exhibit both alignment and misalignment with the foundational principles and properties of language assessment outlined in the literature. We now examine how these principles and properties are reflected in the assessment practices found in the study:

- Practicality involves considerations such as cost, time, ease of administration, scoring, and interpretation (Brown, 2004; Malone, 2011). In the study's context, the assessments, particularly the unit tests derived from the course curriculum, align well with the practicality principle. They are integrated into the course structure and can be administered efficiently within the institution's resource constraints.
- Reliability pertains to the consistency of test results across various factors. It requires clear instructions, uniform rubrics, and unambiguous items or tasks (Bachman & Palmer, 1996; Brown & Abeywickrama, 2010). The study's assessments demonstrate a degree of reliability, especially in terms of uniformity, as they are derived from a standardized curriculum with established content and procedures.

- Validity is often deemed the most critical principle in language assessment. It ensures that a test accurately measures the intended construct (Brown & Abeywickrama, 2010). The alignment between the curriculum objectives, the assessment constructs, and the PET exam's construct validity is a point of concern highlighted in the study. While the curriculum-based assessments align with the curriculum objectives, they may not adequately prepare students for the broader construct of the PET exam. This raises questions about the validity of the assessments in accurately measuring the listening abilities expected at the B1 level.
- Authenticity stresses the importance of assessment tasks mirroring real-world language use (Bachman & Palmer, 1996; Brown & Abeywickrama, 2010). The assessment in the study, particularly the unit tests, demonstrates a degree of authenticity by incorporating tasks such as listening to audio media and recordings. However, the lack of exposure to diverse communicative listening activities, which are considered more authentic, is a limitation.
- Washback refers to the influence of assessments on teaching, learning, motivation, materials, and classroom practices (Hughes, 2003; McKinley & Thompson, 2018). The study suggests that the misalignment between the curriculum-based assessments and the PET exam can lead to a lack of effective washback. Students may not be adequately prepared for the external exam, impacting their motivation and learning outcomes.
- Construct validity emphasizes that a test should measure the intended construct or ability (Bachman & Palmer, 1996). In the study, there is a notable difference in the construct validity of the curriculum-based assessments and the PET exam. The curriculum assessments primarily focus on phonological knowledge, while the PET exam assesses a broader range of listening skills. This misalignment questions the construct validity of the curriculum assessments in preparing students for the PET exam.
- Interactiveness and impact identified by Bachman and Palmer (1996) highlight the need for assessment to be interactive and to have a significant impact on teaching and learning. While the PET exam is designed to be interactive and impact teaching and learning by serving as an external benchmark, the curriculum-based assessment may fall short in terms of interactiveness and impact, particularly if it inadequately prepares students for the external exam.
- Practicality, as described by Bachman and Palmer (1996), encompasses elements of ease of use and usefulness. While the curriculum-based assessments may be practical in terms of their integration into the course, questions arise about their usefulness in adequately preparing students for the PET exam.

In summary, the assessment practices found in the study align with practicality and, to some extent, reliability principles. However, there are notable misalignments with the principles of validity, authenticity, and washback, particularly in the context of preparing students for an external proficiency exam like the PET. This suggests the need for curriculum adjustments and teacher assessment literacy (Giraldo, 2018, 2021a, 2021b; Giraldo et al., 2023; Giraldo & Murcia, 2018, 2019) to address these misalignments and enhance the effectiveness of listening assessment practices.

## Conclusions

This paper has sought to characterize the listening assessment used in B1-level classes at a private language institution in Colombia to understand why the learners' listening skills are consistently below the expected level when taking a proficiency test such as the Preliminary English Test (PET).

The findings discern that this course's pedagogical approach to listening assessment predominantly centers on developing word recognition and decoding skills. As students progress through the course, they are gradually exposed to the broader context of understanding speakers' intent and the dynamics of complete conversations, aligning with the principles articulated in the course objectives. Unit tests from the CCTP predominantly assess phonological knowledge, specifically recognizing spoken words and lexical phrases, chiefly through listening to audio media and recordings. Intriguingly, the midterm exam, also stemming from the CCTP, delves deeper into the listening assessment landscape. It targets the recognition of spoken words and phrases and evaluates the learners' capacity to follow the flow of information within a listening context. These examinations aspire to replicate the characteristics of internationally recognized English language tests, albeit with a notably limited inclusion of only two listening tasks throughout the B1 course. This could leave students ill-prepared for the final exit exam, which is a PET mock assessment.

These findings expose a misalignment between the course listening pedagogy, assessment practices, and the use of an external standardized test as the ultimate exit examination. While the decision to employ the PET exam as the exit assessment is seemingly logical, considering that students complete the B1 curriculum before undertaking it, the analysis has unveiled a discernible discrepancy. Anticipating that students can demonstrate CEFR B1-level listening competencies after engaging in a curriculum primarily grounded in word recognition has proven ineffectual, as substantiated by the consistent underperformance in the listening section.

This study suggests two main implications: curriculum alignment and teachers' assessment literacy. In consonance with Latimer's (2009) recommendations, if the institution elects to

retain the PET mock exam as the B1 level's concluding assessment, it becomes imperative to integrate extracurricular activities dedicated to PET preparation. Such an approach mitigates disruptions to the holistic communicative language mission of the course curriculum. Furthermore, our study resonates with other researchers' advocacy for incorporating specific test-taking skills and familiarity with test items throughout the course. This encompassing approach includes a comprehensive analysis of question types and text genres students can encounter in the PET exam. This alignment should consider the principles and properties of language assessment, ensuring practicality, reliability, validity, authenticity, and washback.

Concerning teachers' assessment literacy, the findings underscore the importance of providing teachers with comprehensive training in listening assessment (Gamboa-Mena & Sevilla-Morales, 2015; Giraldo et al., 2023; Giraldo, 2018, 2021a, 2021b; Giraldo & Murcia, 2018, 2019). Enhancing listening test design practices can be achieved through targeted teacher training, subsequently ameliorating assessment quality. Therefore, it is the institution's imperative to contemplate providing training in listening assessment methods for its teaching faculty.

This study augments the meager corpus of empirical insights regarding listening assessment within the Colombian context. It extends a contextual comprehension of listening and its assessment within the EFL classroom, affording teachers and institutions the capacity to make judicious decisions, adaptations, or proposals harmonious with learners' unique characteristics and exigencies within the prescribed curriculum.

The findings of this study, originating from a specific private language institution in Colombia during the period spanning 2017 to 2019, should be cautiously interpreted within this localized context. While they yield valuable insights into listening assessment practices, the results may differ from the diverse approaches adopted by other institutions or changes that may have occurred since the study's timeframe. Additionally, this study principally relies on quantitative data, with a limited exploration of the nuanced perspectives of both teachers and learners. Consequently, future qualitative research may offer a more comprehensive understanding of the intricate challenges and potential opportunities inherent in listening instruction and assessment practices.

Subsequent research could further expand the scope of this study by undertaking comparative analyses encompassing multiple language institutions across Colombia. Such comparative investigations might reveal commonalities, distinctions, and best practices in listening assessment within the broader Colombian EFL landscape. Moreover, longitudinal studies tracking students' linguistic progression, specifically emphasizing their listening skills, could offer insights into the enduring impact of varied instructional approaches and assessment methodologies. Additionally, research exploring the efficacy of teacher training programs, particularly those that bolster assessment literacy and refine listening instruction,

could shed light on how such training influences classroom practices and, in turn, student outcomes within this EFL context.

## References

Alderson, C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129. https://doi.org/10.1093/applin/14.2.115

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.

Ballesteros-Muñoz, L., & Tutistar-Jojoa, S. (2014). How setting goals enhances learners' self-efficacy beliefs in listening comprehension. *HOW*, 21(1), 42-61. https://doi.org/10.19183/how.21.1.14

Bhattacharya, K. (2009). Othering Research, Researching the Other: De/Colonizing Approaches to Qualitative Inquiry. In J. c. Smart (Eds.), *Higher education: handbook of theory and research. higher education* (Vol. 24, pp, 105-150). Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-9628-0_3

Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. Longman.

Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education.

Buck, G. (2001). *Assessing listening*. Cambridge University Press.

Cárcamo-Morales, B. (2018). Types of listening comprehension developed in the Chilean EFL textbook Global English. *Colombian Applied Linguistics Journal*, 20(1), 49-61. https://doi.org/10.14483/22487085.12313

Córdoba-Zúñiga, E., & Rangel-Gutiérrez, E. (2018). Promoting listening fluency in pre-intermediate EFL learners through meaningful oral tasks. *Profile: Issues in Teachers' Professional Development*, 20(2), 161-177. https://doi.org/10.15446/profile.v20n2.62938

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (1st ed.). Cambridge University Press.

Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal,* 57(4), 325-334.

Gamboa-Mena, R., & Sevilla-Morales, H. (2015). The impact of teacher training on the assessment of listening skills. *LETRAS,* (57), 77-102.

Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development, 20*(1), 179-195. https://doi.org/10.15446/profile.v20n1.62089

Giraldo, F. (2021a). Language assessment literacy and teachers' professional development: A review of the literature. *Profile: Issues in Teachers' Professional Development, 23*(2), 265–279. https://doi.org/10.15446/profile.v23n2.90533

Giraldo, F. (2021b). Language assessment literacy: Insights for educating English language teachers through assessment. *HOW, 28*(3), 78–92. https://doi.org/10.19183/how.28.3.673

Giraldo, F., & Murcia, D. (2018). Language assessment literacy for pre-service teachers: Course expectations from different stakeholders. *GiST: Education and Learning Research Journal*, 16, 56-77. https://doi.org/10.26817/16925777.425

Giraldo, F., & Murcia, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal, 21*(2), pp. 243-259. https://doi.org/10.14483/22487085.14514

Giraldo, F., Escalante-Villa, D., & Isaza-Palacio, D. (2023). English teachers' perceptions of a language assessment literacy course. *HOW, 30*(2), 70–91. https://doi.org/10.19183/how.30.2.750

Green, R. (2019). *Designing Listening Tests*. Palgrave Macmillan.

Hernández-Ocampo, S. P., & Vargas, M. C. (2013). Encouraging students to enhance their listening performance. *Profile Issues in Teachers' Professional Development*, 15(2), 199-214.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.

Islam, M. A., & Aldaihani, F. M. F. (2022). Justification for adopting qualitative research method, research approaches, sampling strategy, sample size, interview method, saturation, and data analysis. *Journal of International Business and Management*, 5(1), 01-11.

Jiménez, M., Rodríguez, C., & Rey Paba, L. (2017). Standardized test results: An opportunity for English program improvement. *HOW*, 24(2), 121-140. http://dx.doi.org/10.19183/how.24.2.335

Latimer, D. G. (2009). *Washback effects of the Cambridge Preliminary English Test at an Argentinean bilingual school* [Doctoral dissertation]. University of Georgia.

Malone, M. (2011). Assessment literacy for language educators. *CALDigest,* October, 1-2.

Mayora, C. A. (2017). Extensive listening in a Colombian university: Process, product, and perceptions. *HOW,* 24(1), 101-121. https://doi.org/10.19183/how.24.1.311

McKinley, J., & Thompson, G. (2018). Washback effect in teaching English as an international language. *The TESOL Encyclopedia of English Language Teaching*, 1–12. https://doi.org/10.1002/9781118784235.eelt0656

Merifield, S., Holmes, L., Walker, S., Wyatt, R., & Walter, C. (2015). *Navigate: Pre-intermediate B1: Teacher's Guide with Teacher's Support and Resource Disc*. Oxford University Press.

Merriam, S. (1988). *Case study research in education: A qualitative approach*. Jossey-Bass.

Morales, A., & Beltrán, A. M. (2006). Developing listening through the use of authentic material. *HOW*, 13(1), 101-124. https://www.howjournalcolombia.org/index.php/how/article/view/110

Mousavi, S. A. (2009). *An encyclopedic dictionary of language testing* (4th ed.). Rahnama Publications.

Preliminary English Test: Handbook. (2005). *University of Cambridge ESOL Examinations.*

Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Longman.

Sevilla-Morales, H., & Chaves-Fernandez, L. (2019). Authentic assessment in the listening comprehension classroom: Benefits and implications. *GiST – Education and Learning Research Journal*, (19), 6-30. https://doi.org/10.26817/16925777.704

Stake, R. E. (2005). Qualitative Case Studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research* (pp. 443–466). Sage Publications Ltd.

Worthington, D. L., & Bodie, G. D. (2017). *The sourcebook of listening research: Methodology and measures*. John Wiley & Sons.

Yin, R. K. (2013). *Case study research: Design and methods* (5th ed.). Sage.

## Appendix A
## Checklist: Type of knowledge in CCTP

| Types of Knowledge | Specific Attributes | TEST 3 Pt 1 | TEST 3 Pt 2 | TEST 4 | TEST 6 | TEST 8 Pt 1 | TEST 8 Pt 2 | TEST 11 Pt 1 | TEST 11 Pt 2 | Exam Pt 1 | Exam Pt 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Phonological | Recognition of phonemes and phonemic clusters | | | | | | | | | | |
| | Knowledge of allophonic variations in fast speech | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| | Knowledge of prosody, intonation, and stress | | | | | | | | | | |
| | Spoken recognition of words and lexical phrases | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Lexical | Recognition of basic word forms | | | | | | | | | | |
| | Knowledge of the meaning of words | | | | | | | | | | |
| | knowing of lexical relationships, collocations, | | | | | | | | | | |
| | Syntactic knowledge (allowable forms of words) | | | | | | | | | | |

| Types of Knowledge | Specific Attributes | TEST 3 Pt 1 | TEST 3 Pt 2 | TEST 4 | TEST 6 | TEST 8 Pt 1 | TEST 8 Pt 2 | TEST 11 Pt 1 | TEST 11 Pt 2 | Exam Pt 1 | Exam Pt 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Syntactic | Ability to perform sentence-level parsing (understand basic grammar within pause unit) | | | | | | | | | | |
| | Ability to perform discourse-level parsing (grammar across pause unit) | | | | | | | | | | |
| | Recognizing collocations | | | | | | | | | | |
| | Understanding cohesion markers in discourse | | | | ✓ | | | | | | |
| Pragmatic | Following the flow of given vs. new information | | | | | | | | | ✓ | ✓ |
| | Inferring speaker intention and motivation | | | | | | | | | | |
| | Recognizing intertextuality (cultural references) | | | | | | | | | | |
| | Understanding social and cultural conventions | | | | | | | | | | |
| | Understanding relationships between interlocutors | | | | | | | | | | |

| Types of Knowledge | Specific Attributes | TEST 3 Pt 1 | TEST 3 Pt 2 | TEST 4 | TEST 6 | TEST 8 Pt 1 | TEST 8 Pt 2 | TEST 11 Pt 1 | TEST 11 Pt 2 | Exam Pt 1 | Exam Pt 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| General | Content/background knowledge | | | | | | | | | | |
| | Extra linguistic knowledge (visual context, gestures, facial expressions) | | | | | | | | | | |
| | Paralinguistic knowledge (prosodic features) | | | | | | | | | | |
| | Social and pragmatic knowledge | | | | | | | | | | |
| | Strategic knowledge (social, cognitive, and metacognitive strategies) | | | | | | | | | | |

## Appendix B
## Checklist: Type of knowledge in ETP

| Types of Knowledge | Specific Attributes | PET Part 1 | PET Part 2 | PET Part 3 | PET Part 4 |
|---|---|---|---|---|---|
| Phonological | Recognition of phonemes and phonemic clusters | | | | |
| | Knowledge of allophonic variations in fast speech | | | | |
| | Knowledge of prosody, intonation, and stress | | | | |
| | Spoken recognition of words and lexical phrases | | | ✓ | |
| Lexical | Recognition of basic word forms | ✓ | ✓ | ✓ | |
| | Knowledge of the meaning of words | | | | |
| | knowing of lexical relationships, collocations, | | | | |
| | Syntactic knowledge (allowable forms of words) | | | | |
| Syntactic | Ability to perform sentence-level parsing (understand basic grammar within pause unit) | | | | |
| | Ability to perform discourse-level parsing (grammar across pause unit) | | | | |
| | Recognizing collocations | | | | |
| | Understanding cohesion markers in discourse | | | | |
| Pragmatic | Following the flow of given vs. new information | ✓ | ✓ | ✓ | ✓ |
| | Inferring speaker intention and motivation | | | | |
| | Recognizing intertextuality (cultural references) | | | | |
| | Understanding social and cultural conventions | | | | |
| | Understanding relationships between interlocutors | | | | |
| General | Content/background knowledge | | | | |
| | Extra linguistic knowledge (visual context, gestures, facial expressions) | ✓ | | | |
| | Paralinguistic knowledge (prosodic features) | | | | |
| | Social and pragmatic knowledge | | | | |
| | Strategic knowledge (social, cognitive, and metacognitive strategies) | | | | |

## Appendix C
## Checklist: Forms of Assessment in CCTP

| Type of test | Type of Task | TEST 3 Pt 1 | TEST 3 Pt 2 | TEST 4 | TEST 6 | TEST 8 Pt 1 | TEST 8 Pt 2 | TEST 11 Pt 1 | TEST 11 Pt 2 | Exam Pt 1 | Exam Pt 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Discrete Item | Multiple choice questions | | | | | | | | | | |
| | Open questions | | | | | | | | | ✓ | |
| Task-Based | Closed task involving single response | | | | | | | | | | ✓ |
| | Open tasks involving multiple responses | | | | | | | | | | |
| Integrative | Memory test following or during listening to an extract. | | | | | | | | | | |
| | Dictation, complete or partial | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| Communicative | Written communicative tasks involving listening (such as writing a complaint letter after hearing a problem description). | | | | | | | | | | |
| | Oral tasks involving listening (such as following directions on a map). | | | | | | | | | | |
| Interview | Face-to-face performances with the teacher or another student | | | | | | | | | | |
| | Extended oral interviews | | | | | | | | | | |

## Appendix D
## Checklist: Forms of Assessment in ETP

| Type of test | Type of Task | PET Part 1 | PET Part 2 | PET Part 3 | PET Part 4 |
|---|---|---|---|---|---|
| Discrete Item | Multiple choice questions | ✓ | ✓ | | |
| | Open questions | | | | ✓ |
| Task-Based t | Closed task involving single response | | | ✓ | |
| | Open tasks involving multiple responses | | | | |
| Integrative | Memory test following or during listening to an extract. | | | | |
| | Dictation, complete or partial | | | | |
| Commu-nicative | Written communicative tasks involving listening (such as writing a complaint letter after hearing a description of a problem). | | | | |
| | Oral tasks involving listening (such as following directions on a map). | | | | |
| Interview | Face-to-face performances with the teacher or another student | | | | |
| | Extended oral interviews | | | | |

## Appendix E
## Examples of Can-Do Statements at the B1 level - CEFR

Examples of Can Do statements at Level B1

| Typical abilities | Listening and Speaking |
|---|---|
| **Overall general ability** | CAN understand straightforward instructions or public announcements.<br>CAN express simple opinions on abstract/cultural matters in a limited way. |
| **Social & Tourist** | CAN identify the main topic of a news broadcast on TV if there is a strong visual element.<br>CAN ask for information about accommodation and travel. |
| **Work** | CAN follow a simple presentation/demonstration.<br>CAN offer advice to clients within own job area on simple matters. |
| **Study** | CAN understand instructions on classes and assignments given by a teacher or lecturer.<br>CAN take part in a seminar or tutorial using simple language. |

## Appendix F
## PET Listening Section Results (2017, 2018, and 2019)