AN EVALUATION FRAMEWORK FOR SECOND LANGUAGE LISTENING COURSEBOOKS: PRINCIPLES AND PRACTICE

Jiayu Zhai[a, b], Icy Lee[b], Vahid Aryadoust[b*], Willy Ardian Renandya[b]
([a]nie23.zj0545@e.ntu.edu.sg; [b]icy.lee@nie.edu.sg; [b]vahid.aryadoust@nie.edu.sg; [b]willy.renandya@nie.edu.sg)

[a] *School of English Studies, Sichuan International Studies University, No. 33, Zhuangzhi Road, Shapingba District, Chongqing, China*

[b] *National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore*

Abstract: Coursebooks are central to second language (L2) listening instruction, yet existing evaluation tools tend to be generic rather than specifically tailored to L2 listening. To offer teachers principled guidance for selecting and evaluating instructional materials, this paper proposes an evaluation framework for L2 listening coursebooks, grounded in theories of listening comprehension, listening assessment, and second language acquisition. The framework comprises seven principles, each operationalized through a guiding question and a set of indicators. An illustrative example is also provided by applying the framework to *Listen This Way* (Level 4), a coursebook used in undergraduate English programs in the Chinese mainland, demonstrating how the framework supports a systematic evaluation of an L2 listening coursebook. Implications for L2 listening instruction and coursebook design are discussed.

Keywords: coursebook evaluation, listening coursebooks, listening tasks, listening texts, second language listening

DOI: https://doi.org/10.15639/teflinjournal.v37i1/210-233

Listening is a fundamental skill in second language (L2) acquisition, providing learners with language input and supporting overall language development (Renandya & Hu, 2018; Richards, 2005). It also constitutes a core component of interactional competence, essential for effective communication (Galaczi & Taylor, 2018). Given its importance, listening courses are typically compulsory for English majors, particularly in contexts where English is taught as a foreign language (EFL). In such settings, the choice of an L2 listening coursebook is consequential. For teachers, the coursebook serves as the pedagogical tool that shapes instructional approaches and, in turn, learning outcomes (Hutchinson & Torres, 1994; Richards, 2014). For students, especially EFL learners with limited exposure to authentic English use, the coursebook provides essential sources of language input that matter to the development of listening proficiency (Hill & Tomlinson, 2013; Jones, 2022).

The existing literature offers a range of frameworks and checklists for evaluating language-teaching coursebooks (e.g., Cunningsworth, 1995; Ellis, 1997; Littlejohn, 2011; McDonough et al., 2013; Tomlinson, 2013; see Mukundan & Ahour, 2011, for a review of textbook evaluation

checklists). These frameworks typically address general dimensions of coursebook design (e.g., aims, syllabus organization, and content) and cover the four language skills where listening is treated as one component among several. While valuable, this generic, multi-skill approach provides limited guidance on dimensions central to L2 listening. For example, existing frameworks rarely address in detail whether the listening texts in coursebooks reflect the linguistic and discoursal features of authentic spoken language, expose learners to the range of English accents encountered in real-world communication, or incorporate the multimodal characteristics of authentic listening contexts. With respect to listening tasks, they likewise offer limited guidance on whether tasks resemble real-world listening scenarios or engage learners in the cognitive and metacognitive processes that effective listening requires. Such dimensions pertaining to listening texts and tasks are foregrounded in research on listening assessment, listening comprehension, and second language acquisition, but fall outside the scope of general coursebook evaluation tools. The present paper addresses this gap by proposing an evaluation framework focused specifically on L2 listening, extending broader coursebook evaluation approaches into this underexplored domain.

The development of the proposed framework was guided by three interrelated questions about listening texts and tasks: (1) What good qualities should listening texts and tasks exhibit? (2) What cognitive processes should listening tasks engage learners in? (3) What level of difficulty should listening texts present to support learning? The first question draws on theories of listening assessment, particularly the concepts of authenticity and cognitive validity (e.g., Bachman & Palmer, 1996; Buck, 2001; Field, 2013; Ockey & Wagner, 2018; Weir, 2005). The second draws on theories of listening comprehension, which specify the cognitive and metacognitive processes underlying effective listening (e.g., Anderson, 1995; Field, 2004, 2008, 2013; Goh & Vandergrift, 2022). The third draws on second language acquisition, particularly the comprehensible input hypothesis (Krashen, 1981). Together, these three streams of research serve as the conceptual foundation for the framework.

Three criteria also guide the formulation of the principles in this framework. First, each principle is theoretically grounded in one or more of the three research streams and reflects a well-established construct in that literature. Second, each principle is listening-specific, addressing dimensions central to L2 listening. Third, each principle is operationalizable through observable features of listening materials and tasks.

Informed by the recurring constructs in listening assessment, listening comprehension, and second language acquisition relevant to L2 listening coursebook evaluation, seven principles were formulated, as presented in Table 1.

Table 1. Seven Principles for Evaluating L2 Listening Coursebooks

Principles
1. Listening texts should reflect typical linguistic features of authentic spoken language.
2. Listening texts should include diverse English accents.
3. Listening texts should be multimodal.

Principles
4. The difficulty of listening texts should align with learners' proficiency level.
5. Listening tasks should be varied and resemble real-world listening scenarios.
6. Listening tasks should engage learners in both bottom-up and top-down processing.
7. Listening tasks should facilitate the development of metacognitive awareness.

These principles address two core components of L2 listening coursebooks: the selection of listening texts (Principles 1 to 4) and the design of listening tasks (Principles 5 to 7). In terms of theoretical grounding, Principles 1, 2, 3, and 5 draw on theories of language assessment and empirical studies addressing the authenticity and validity of listening texts and tasks (Aryadoust, 2024; Aryadoust & Luo, 2023; Bachman & Palmer, 1996; Batty & Suvorov, 2024; Buck, 2001; Field, 2013; Ockey & Wagner, 2018; Taylor & Geranpayeh, 2011; Wagner, 2018; Weir, 2005). Principles 6 and 7 are grounded in theories of listening comprehension, pertaining to the cognitive and metacognitive processes that listening tasks should engage learners in (Anderson, 1995; Field, 2004, 2013; Goh & Vandergrift, 2022). Principle 4 draws on theories of second language acquisition, concerning how the difficulty of listening texts should align with learner proficiency (Krashen, 1981). Table 2 summarizes the connection between the three theoretical streams, their associated constructs, and the corresponding principles. These principles are distributed unevenly across the three theoretical streams because they were derived from constructs relevant to L2 listening coursebook evaluation, rather than to give these streams equal weight. Other constructs relevant to L2 listening, such as affective and sociocultural dimensions, were considered but excluded, as they depend more on learner characteristics or classroom interaction than on observable features of coursebook materials, thereby making them less amenable to coursebook evaluation.

Table 2. Theoretical Streams, Key Constructs, and Corresponding Principles

Theoretical Stream	**Associated Construct(s)**	**Principle(s)**
Listening assessment (Aryadoust, 2024; Aryadoust & Luo, 2023; Bachman & Palmer, 1996; Batty & Suvorov, 2024; Buck, 2001; Field, 2013; Ockey & Wagner, 2018; Taylor & Geranpayeh, 2011; Wagner, 2018; Weir, 2005)	Authenticity; cognitive validity	1, 2, 3, 5
Listening comprehension (Anderson, 1995; Field, 2004, 2008, 2013; Goh & Vandergrift, 2022)	Cognitive processes; metacognitive awareness	6, 7
Second language acquisition (Krashen, 1981)	Comprehensible input ($i + 1$)	4

Note. The framework integrates three theoretical streams and their associated constructs to derive seven principles for evaluating L2 listening coursebooks.

AN EVALUATION FRAMEWORK FOR L2 LISTENING COURSEBOOKS

This section discusses each of the seven principles in turn. For each principle, we present its theoretical grounding, a guiding question that captures its core concern, and a set of indicators through which teachers can operationalize it in coursebook evaluation.

Principle 1: Listening texts should reflect typical linguistic features of authentic spoken language

Principle 1 is grounded in the concept of authenticity in listening assessment. This principle concerns the linguistic features of listening texts in coursebooks and emphasizes the importance of exposing learners to spoken discourse that reflects real-world language use. Overly scripted listening texts have long been criticized for underrepresenting features of spoken language and leaving learners underprepared for authentic communication (Wagner, 2014; Wagner & Toth, 2014). This concern has also been documented in coursebook-specific analyses, in which dialogues have been shown to systematically lack the discourse features of authentic interaction (Gilmore, 2004). To better prepare L2 learners for real-world listening, listening texts should incorporate linguistic features commonly found in authentic spoken communication (Ockey & Wagner, 2018; Wagner, 2018).

Research has identified a range of linguistic features associated with authentic spoken language (e.g., Biber, 1988; Brown & Crowther, 2023; McCarthy, 2010; McCarthy & Carter, 1995; Wagner & Toth, 2014). These include (1) phonological features, such as contraction, reduction, linking, elision, and assimilation, which contribute to the natural flow of connected speech; (2) grammatical features, such as shorter idea units, fewer embedded clauses, reduced nominalization, and generally less dense syntactic structure; (3) lexical features, such as informal and conversational expressions commonly used in spoken interaction; (4) discourse features, such as circumlocution, false starts, hesitation phenomena, fillers, repetitions, and self-repair; and (5) interactive features, such as turn-taking, turn-openers, and backchannels.

To operationalize this principle, teachers may first form a holistic judgment by asking the following guiding question: *Do the listening texts sound like natural spoken discourse rather than written scripts read aloud?* They may then draw on more specific indicators to support this judgment: (1) the spoken delivery contains features of connected speech, such as reduction, linking, elision, or assimilation; (2) the language is organized in relatively short idea units rather than densely packed, written-style syntax; (3) the text includes conversational or informal lexical expressions where appropriate; (4) the discourse contains features such as fillers, hesitation, repetition, or self-repair; and (5) in interactive texts, the exchange includes features such as turn-taking and backchanneling.

Principle 2: Listening texts should include diverse English accents

Principle 2 is also informed by the concept of authenticity in listening assessment. This principle concerns whether listening texts reflect the diversity of English accents that learners are likely to encounter in real-world communication. As the use of English has expanded beyond traditionally English-speaking countries to what Kachru (1985) described as the outer circle, where English functions as a second language (e.g., India), and the expanding circle, where it is learned as a foreign language (e.g., China and Japan), English now operates widely as a lingua franca among speakers from diverse linguistic backgrounds (Matsuda, 2019). Accordingly, L2 learners are likely to encounter a wide range of English accents in real-world communication.

For this reason, listening texts should not be confined to the British and American accents that predominate in L2 listening coursebooks. Instead, a broader range of English varieties should be incorporated into coursebooks. Doing so not only enhances the authenticity of listening texts but also better prepares learners to communicate in English across linguistically and culturally diverse contexts (Aryadoust, 2024; Harding, 2012; Kang et al., 2024; Karakaş, 2024).

To operationalize this principle, teachers may ask the guiding question: *Do the listening texts include a range of English accents rather than British and American accents alone?* More specific indicators include: (1) the presence of inner-circle varieties beyond British and American English, such as Australian or Canadian English; (2) the presence of outer-circle varieties, such as Indian or Malaysian English; and (3) the presence of expanding-circle varieties, such as Chinese- or Japanese-accented English.

Principle 3: Listening texts should be multimodal

Grounded in the concepts of authenticity and construct validity in listening assessment, Principle 3 advocates the use of multimodal listening texts in coursebooks to reflect the nature of real-world listening and to adequately represent the construct of L2 listening. Listening is rarely a purely auditory activity. In academic and non-academic contexts, such as lectures, classroom discussions, and everyday conversation, listeners often process auditory and visual information simultaneously (Kwon & Yu, 2024; Ockey & Wagner, 2018; Taylor & Geranpayeh, 2011). Visual information typically includes visual aids, such as slides, pictures, and written texts, as well as nonverbal cues, such as gestures and facial expressions (Batty & Suvorov, 2024). These visuals can significantly affect listening comprehension (Kwon & Yu, 2024; Sueyoshi & Hardison, 2005). In response, researchers have called for a broader conceptualization of L2 listening that recognizes visual processing as an integral component of listening ability (e.g., Aryadoust & Luo, 2023; Ockey & Wagner, 2018; Wagner, 2008, 2010).

Listening texts in coursebooks should not rely on audio alone but incorporate visual information. Such visuals can be static (e.g., pictures, diagrams, and written texts) or dynamic (e.g., videos and animations), and can combine different types of visual information, including visual aids and speaker-related nonverbal cues. By exposing learners to multimodal input, L2 listening coursebooks can better prepare them for real-world communication where meaning-making often involves integrating both auditory and visual information.

To operationalize this principle, teachers may ask this guiding question: *Do the listening texts combine auditory input with relevant visuals?* More specific indicators include: (1) the presence of static visuals such as pictures, charts, diagrams, and other visual aids; (2) the presence of dynamic visuals such as videos or animations; and (3) the inclusion of speaker-related nonverbal cues such as gestures and facial expressions.

Principle 4: The difficulty of listening texts should align with learners' proficiency level

Principle 4 is informed by the notion of comprehensible input in second language acquisition, particularly Krashen's (1981) concept of $i + 1$, which holds that input should be slightly beyond learners' current competence in order to drive language development. Texts that are too easy offer limited opportunities for improvement, whereas texts that are too difficult risk overloading learners, impeding comprehension, and reducing engagement. Listening texts in coursebooks should therefore be broadly understandable while still presenting an appropriate degree of challenge. Aligning text difficulty with learner proficiency in this way supports the gradual development of listening ability.

To evaluate the difficulty of listening texts, teachers may draw on three complementary sources of evidence. First, they can consider the coursebook's stated target proficiency level as the initial reference, such as the intended learner group or CEFR band. Second, they can use their professional judgment, drawing on their familiarity with students' listening ability and learning needs, to determine whether the texts are comprehensible yet appropriately challenging. Third, they can consult analytical tools such as Text Inspector, a web-based application that analyzes linguistic features and provides measures of vocabulary profile, complexity, and estimated CEFR level (Bax, 2012). Used together, these sources can help teachers make an informed judgment about whether the difficulty of listening texts is appropriate for their students.

To operationalize this principle, teachers can ask the guiding question: *Are the listening texts comprehensible yet slightly challenging for the students?* More specific indicators include: (1) the alignment between the stated target proficiency level of the coursebook and students' actual proficiency level; (2) teachers' professional judgment that the listening texts are comprehensible while still providing a modest degree of challenge; and (3) where available, difficulty estimates from analytical tools such as Text Inspector that support the judgment that the texts are neither too easy nor too difficult.

Principle 5: Listening tasks should be varied and resemble real-world listening scenarios

Principle 5 is informed by the concepts of authenticity and cognitive validity in language assessment. Authenticity suggests that listening tasks should resemble real-world tasks (Bachman & Palmer, 1996; Buck, 2001), while cognitive validity further requires that such tasks should engage cognitive processes similar to those involved in real-world listening (Field, 2013; Weir, 2005). Incorporating varied task types is also recommended because different tasks tend to elicit different cognitive activities and strategy use (Alderson, 2000). Listening tasks are broadly classified as integrated and non-integrated tasks (Ockey & Wagner, 2018; Rukthong, 2021). Integrated tasks require learners to use listening together with other language skills, typically by incorporating information from the listening input into an oral or written response,

as in listen-to-speak and listen-to-write tasks (Ockey & Wagner, 2018). Such tasks reflect the complex and interactive nature of real-world listening situations (Rukthong & Brunfaut, 2020). Non-integrated, or independent, tasks focus on listening skills alone and often rely on discrete-point items such as multiple-choice questions, cloze tests, and matching items. Non-integrated tasks can be useful for targeting specific listening subskills, but they may not fully capture the cognitive processes involved in authentic communication contexts where listening is often integrated with speaking and/or writing (Rukthong, 2021). Taken together, these considerations suggest that L2 listening coursebooks should include a variety of task types and maintain a reasonable balance between non-integrated and integrated tasks, supporting both the development of specific listening subskills and learners' preparation for real-world communication.

To operationalize this principle, teachers can ask the guiding question: *Does the coursebook include varied listening task types that resemble real-world listening?* More specific indicators include: (1) the presence of multiple non-integrated task types targeting different aspects of listening comprehension, such as dictation, blank-filling, multiple-choice questions, true/false items, and matching tasks; (2) the inclusion of integrated tasks such as listen-to-speak and listen-to-write; and (3) a reasonable balance between non-integrated and integrated tasks across the coursebook.

Principle 6: Listening tasks should engage learners in both bottom-up and top-down processing

Principle 6 is informed by theories of listening comprehension, which identify bottom-up and top-down processing as two essential cognitive processes (Anderson, 1995; Field, 2004, 2013; Goh & Vandergrift, 2022). Bottom-up processing involves decoding lower-level linguistic units, such as sounds, words, and clauses, in order to construct meaning, whereas top-down processing involves drawing on prior knowledge, contextual information, and discourse expectations to interpret and infer meaning (Field, 2004). These two processes interact throughout the three phases of listening comprehension: perception, parsing, and utilization (Anderson, 1995; Goh & Vandergrift, 2022). During perception, listeners primarily engage in bottom-up processing to recognize basic linguistic input. During parsing, bottom-up and top-down processing work together as listeners segment the speech stream using syntactic and semantic cues and organize information into meaningful propositions. During utilization, listeners rely more heavily on top-down processing to integrate contextual knowledge and prior schemata with linguistic input, ultimately constructing a coherent mental representation of the listening text.

Different listening tasks can be used to engage and develop bottom-up and top-down processing. Tasks focusing on recognizing words, identifying details, and extracting explicit information typically target bottom-up processing, whereas tasks requiring learners to infer meaning, identify main ideas, and recognize discourse organization often target top-down processing (e.g., Cubillos et al., 2008; Field, 2004; Yeldham & Gruba, 2014). A well-designed coursebook should include listening tasks that engage learners in both processes.

To operationalize this principle, teachers can ask the guiding question: *Do the listening tasks engage students in both bottom-up and top-down processing?* More specific indicators include: (1) the presence of tasks that require learners to identify sounds, words, key details, and explicit information (e.g., dictation, blank-filling, and detail-oriented questions); and (2) the presence of tasks that require learners to infer meaning, identify main ideas, interpret speaker intention, or recognize discourse structure (e.g., inferential questions, main-idea questions, and open-ended questions with "why" or "how" prompts).

Principle 7: Listening tasks should facilitate the development of metacognitive awareness.

Principle 7 is informed by the well-established role of metacognitive awareness in effective listening comprehension and in the long-term development of listening ability (Goh & Vandergrift, 2022). This principle highlights the importance of designing coursebook tasks that help students become more aware of how listening works and how they can regulate their own listening processes. In L2 listening, metacognitive awareness comprises two key dimensions: metacognitive knowledge and strategy use (Goh & Vandergrift, 2022). Metacognitive knowledge consists of person knowledge, task knowledge, and strategy knowledge. Person knowledge refers to learners' awareness of their own strengths, weaknesses, and affective states as listeners; task knowledge refers to their understanding of the nature and demands of different listening tasks; strategy knowledge refers to their understanding of the strategies available for supporting comprehension and addressing problems during listening (Goh & Vandergrift, 2022). Strategy use refers to the actual application of such strategies to plan, monitor, and evaluate listening (Goh & Vandergrift, 2022).

To support the development of metacognitive awareness, listening tasks in coursebooks should be structured to engage these dimensions in a systematic way. Goh and Vandergrift (2022) proposed a range of activities that support the comprehensive development of metacognitive awareness. Such tasks encourage students to set goals, anticipate task demands, monitor their comprehension, reflect on difficulties encountered, and evaluate their listening performance across the pre-, while-, and post-listening phases, offering useful guidance for the design of metacognitive tasks in L2 listening coursebooks. These tasks include: (1) integrated experiential listening tasks, such as the metacognitive pedagogical sequence, which guides students throughout listening stages; (2) self-directed listening activities, which provide prompts for pre-listening preparations and post-listening reflection; (3) guided reflection activities, such as listening diaries with guided questions and process-oriented discussions that encourage learners to evaluate their strategy use; and (4) self-assessment checklists that help them reflect on their knowledge and performance.

To operationalize this principle, teachers can ask the guiding question: *Do the listening tasks help learners plan, monitor, and evaluate their listening?* More specific indicators include: (1) the presence of tasks that prompt learners to reflect on their strengths, weaknesses, or feelings as listeners; (2) the presence of tasks that help learners understand the characteristics and demands of listening tasks; (3) the presence of tasks that introduce useful listening strategies; (4) the presence of tasks that guide students to apply such strategies during listening; and (5) the inclusion of these tasks across pre-listening, while-listening, and post-listening phases.

In sum, this section introduced the seven principles that constitute the proposed framework for evaluating L2 listening coursebooks. Each principle is grounded in relevant theory and research and is operationalized through a guiding question and a set of indicators to support its application. Table 3 summarizes the seven principles alongside their corresponding guiding questions and indicators. In the next section, we present an illustrative example of how the framework can be applied in practice through the evaluation of an L2 listening coursebook.

Table 3. Principles, Guiding Questions, and Indicators of the Evaluation Framework for L2 Listening Coursebooks

Principle	Guiding Question	Indicators
Principle 1 Listening texts should reflect typical linguistic features of authentic spoken language.	Do the listening texts sound like natural spoken discourse rather than written scripts read aloud?	• The spoken delivery contains features of connected speech, such as reduction, linking, elision, or assimilation • The language is organized in relatively short idea units rather than densely packed written-style syntax • The text includes conversational or informal lexical expressions where appropriate • The discourse contains features such as fillers, hesitation, repetition, or self-repair • In interactive texts, the exchange includes features such as turn-taking and backchanneling
Principle 2 Listening texts should include diverse English accents.	Do the listening texts include a range of English accents rather than British and American accents alone?	• The presence of inner-circle varieties beyond British and American English (e.g., Australian or Canadian English) • The presence of outer-circle varieties (e.g., Indian or Malaysian English) • The presence of expanding-circle varieties (e.g., Chinese- or Japanese-accented English)
Principle 3 Listening texts should be multimodal.	Do the listening texts combine auditory input with relevant visuals?	• The presence of static visuals such as pictures, charts, diagrams, and other visual aids • The presence of dynamic visuals such as videos or animations • The inclusion of speaker-related nonverbal cues such as gestures and facial expressions
Principle 4 The difficulty of listening texts should	Are the listening texts comprehensible yet	• The stated target proficiency level of the coursebook aligns with students' actual proficiency level

Principle	**Guiding Question**	**Indicators**
align with learners' proficiency level.	slightly challenging for the students?	• Based on teachers' professional judgment, the listening texts are appropriate for students while still providing a modest degree of challenge • Where available, difficulty estimates from analytical tools such as Text Inspector support the judgment that the texts are neither too easy nor too difficult
Principle 5 Listening tasks should be varied and resemble real-world listening scenarios.	Does the coursebook include varied listening task types that resemble real-world listening?	• The presence of multiple non-integrated task types targeting different aspects of listening comprehension (e.g., dictation, blank-filling, multiple-choice questions, true/false items, and matching tasks) • The inclusion of integrated tasks such as listen-to-speak and listen-to-write • The relative balance between non-integrated and integrated tasks across the coursebook
Principle 6 Listening tasks should engage learners in both bottom-up and top-down processing.	Do the listening tasks engage students in both bottom-up and top-down processing?	• The presence of tasks that require learners to identify sounds, words, key details, and explicit information (e.g., dictation, blank-filling, and detail-oriented questions) • The presence of tasks that require learners to infer meaning, identify main ideas, interpret speaker intention, or recognize discourse structure (e.g., inferential questions, main-idea questions, open-ended questions with "why" or "how" prompts).
Principle 7 Listening tasks should facilitate the development of metacognitive awareness.	Do the listening tasks help learners plan, monitor, and evaluate their listening?	• The presence of tasks that prompt students to reflect on their strengths, weaknesses, or feelings as listeners • The presence of tasks that help students understand the characteristics and demands of listening tasks • The presence of tasks that introduce useful listening strategies • The presence of tasks that guide students to apply such strategies during listening • The inclusion of these tasks across pre-listening, while-listening, and post-listening phases

AN ILLUSTRATIVE EXAMPLE: APPLYING THE FRAMEWORK TO AN L2 LISTENING COURSEBOOK

To demonstrate how the seven principles can be applied in the systematic evaluation of L2 listening coursebooks, this section presents an illustrative application of the framework to evaluate *Listen This Way* (Level 4), a listening coursebook used in undergraduate English programs in the Chinese mainland. The coursebook comprises 11 units, each organized into several sections, including vocabulary preparation, sentence and dialogue listening, skill-focused listening activities, extended input, review tasks, and video-based materials. This coursebook is used with second-year English majors at the university where the first author teaches. The English proficiency of the learner group is approximately intermediate to upper-intermediate, corresponding to B1-B2 on the CEFR. For illustrative purposes, the analysis is limited to Units 1 to 4 of the coursebook.

Principle 1: Listening texts should reflect typical linguistic features of authentic spoken language

Following Principle 1 and its indicators, the linguistic features of listening texts were analyzed. Phonologically, connected-speech phenomena such as linking, reduction, and elision were audibly present. Grammatically, unembedded sentences predominated, accounting for 75.00% of sentences in conversations and 70.65% in mini-lectures, and instances of "spoken grammar" with informal structures (e.g., "I see no reason why not", "You go somewhere else?", and "Why not?") were also evident. Lexically, informal items (e.g., "gonna", "gotta", and "Right?") recurred in conversational exchanges. In terms of discourse features, the texts contained fillers (e.g., "er", "uh") and false starts (e.g., "I'm... I want some money."), along with interactive features such as turn-taking and back-channeling (e.g., "Yeah", "You know"). Together, these observations in the four units of the coursebook support a positive answer to the guiding question: the listening texts broadly resemble natural spoken discourse rather than written scripts read aloud.

The presence of these features suggests that the coursebook generally reflects the emphasis on authenticity that listening assessment research has long advocated (e.g., Buck, 2001; Ockey & Wagner, 2018; Wagner, 2014). Nevertheless, because the texts remain scripted at the production level, they may still lack the inherent spontaneity of unscripted discourse, a limitation that could be mitigated by incorporating unscripted recordings. Pedagogically, teachers can use the listening texts in the coursebook to draw learners' attention explicitly to features of real-world spoken language, while supplementing them with unscripted materials, such as podcasts or interview excerpts, to expose learners to fully spontaneous discourse.

Principle 2: Listening texts should include diverse English accents

The English accents in the listening texts of Units 1 to 4 were analyzed following Principle 2 and its indicators. It was found that the listening texts drew exclusively on two inner-circle varieties, British and American English, with American English predominating (see Figure 1). Other inner-circle varieties (e.g., Australian, Canadian English) were absent, as were outer-circle and expanding-circle varieties. The answer to the guiding question, based on the four-unit

analysis, is therefore negative: the listening texts do not include a range of English accents beyond British and American English.

This finding aligns with prior coursebook research that has identified an over-reliance on British and American English and limited representation of other inner-, outer-, and expanding-circle Englishes in ELT materials (Syrbe & Rose, 2018; Vettorel & Lopriore, 2013). Such a narrow profile of English varieties may inadequately prepare learners for the linguistic diversity of real-world communication, particularly in contexts where English is used as a lingua franca (Aryadoust, 2024; Galloway & Rose, 2018). Pedagogically, teachers can supplement the listening texts with materials featuring a wider range of English accents, for example, additional inner-circle varieties such as Australian English, alongside outer- and expanding-circle varieties such as Indian or Japanese-accented English.

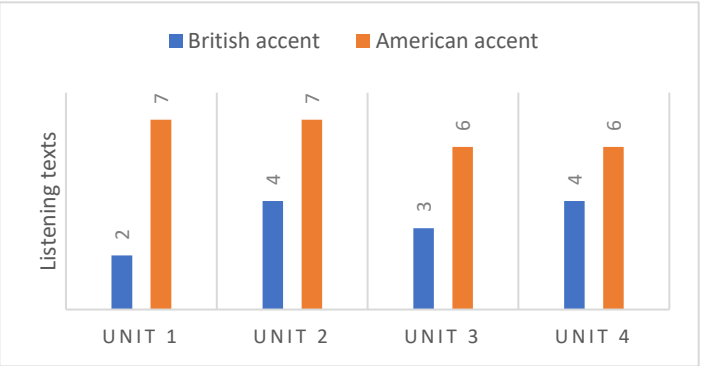


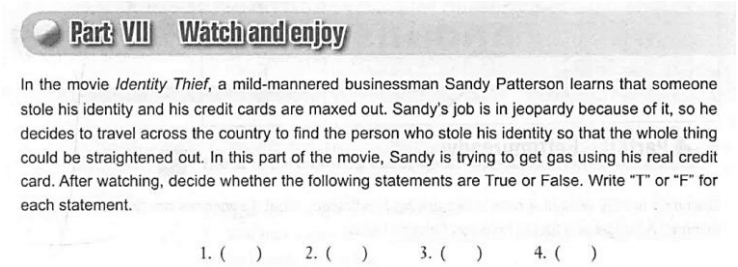
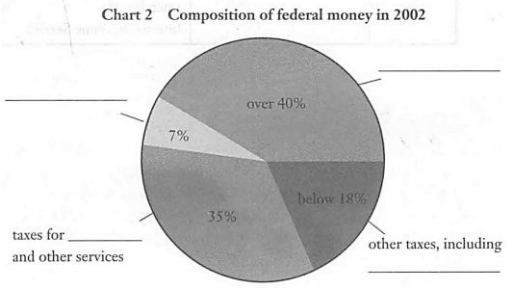
Figure 1. Distribution of British and American Accents in *Listen This Way* (Level 4) (Unit 1-4)

Principle 3: Listening texts should be multimodal

Following Principle 3 and its indicators, the listening texts were analyzed for the presence of multimodality. Dynamic visuals were found to be present in Part 7 of each of the four units, which featured video excerpts from films and carried speaker-related nonverbal cues such as gestures and facial expressions. Static visuals were largely absent, with one exception being Task B in Part 5 of Unit 4, which paired the audio with a pie chart and required students to integrate auditory and visual information (see Table 4). The remaining parts of each unit relied on audio-only input. The answer to the guiding question is therefore mixed: the listening texts do combine auditory input with relevant visuals, but only in limited portions of the coursebook.

The inclusion of multimodal input aligns with the calls in the listening assessment literature to treat L2 listening as a multimodal construct (Batty & Suvorov, 2024; Ockey & Wagner, 2018; Suvorov, 2015; Wagner, 2008, 2010). However, the limited distribution of multimodal texts in the first four units of the coursebook might indicate that multimodality is only treated as an occasional feature rather than a defining characteristic of the listening texts. Pedagogically, teachers can supplement the audio-only materials with more relevant visual aids or draw on additional video sources to expose learners to the kinds of multimodal input they encounter in authentic academic and everyday listening contexts.

Table 4. Examples of Visuals in *Listen This Way* (Level 4)

Visual Type	Example
Dynamic visuals	 Part VII Watch and enjoy In the movie *Identity Thief*, a mild-mannered businessman Sandy Patterson learns that someone stole his identity and his credit cards are maxed out. Sandy's job is in jeopardy because of it, so he decides to travel across the country to find the person who stole his identity so that the whole thing could be straightened out. In this part of the movie, Sandy is trying to get gas using his real credit card. After watching, decide whether the following statements are True or False. Write "T" or "F" for each statement. 1. ( ) 2. ( ) 3. ( ) 4. ( )
Static visuals	

Note. The task examples in this table are from *Listen This Way* (Level 4) (Higher Education Press, China).

Principle 4: The difficulty of listening texts should align with learners' proficiency level

Following Principle 4 and its indicators, the difficulty of the listening texts was examined against students' proficiency level by using the three evidence sources specified in the framework. First, the coursebook's stated target learners are undergraduate English majors, which provides an initial, external reference. Second, based on the first author's teaching experience with the learner group, the listening texts appear broadly appropriate while offering a modest degree of challenge. Third, the listening texts of the four units were analyzed using Text Inspector. On the lexical profile index, where 100% indicates a high level of native-speaker academic text, Units 1 to 4 scored 58.82%, 56.47%, 58.33%, and 60.39%, respectively. The estimated CEFR levels for the four units were C1, B2+, C1, and C1, respectively. Given that the students are undergraduate English majors with intermediate to upper-intermediate language proficiency, these three sources of evidence support an affirmative answer to the guiding question: the listening texts fit the target learner group, remaining comprehensible while presenting a productive degree of challenge.

This alignment between text difficulty and learner proficiency is consistent with Krashen's (1981) comprehensible input hypothesis, which holds that input slightly beyond learners' current level ($i + 1$) supports language development. Pedagogically, teachers can use these listening texts as core material while adjusting task demands to calibrate the level of challenge for individual classes.

Principle 5: Listening tasks should be varied and resemble real-world listening scenarios.

Following Principle 5 and its indicators, the listening task types in the coursebook were analyzed. The coursebook incorporated multiple non-integrated task types, including dictation, blank-filling, open-ended questions, true/false questions, matching, and multiple-choice questions (see Table 5). However, integrated tasks such as listen-to-speak and listen-to-write were absent, leaving the task repertoire entirely within the non-integrated format. The answer to the guiding question is therefore mixed: the four units of the coursebook offer a range of listening task types, but this variety is confined to non-integrated formats and offers limited resemblance to real-world listening.

It should be noted that the variety of non-integrated task types in the coursebook has pedagogical value, as different task types often target different listening subskills and help sustain learner engagement. Nevertheless, the exclusive reliance on non-integrated tasks in the four units of the coursebook contrasts with the advocacy in the listening assessment literature for greater use of integrated tasks, which better reflect the cognitive demands of real-world listening and tap aspects of listening ability that discrete-point items do not measure (Ockey & Wagner, 2018; Rukthong, 2021; Rukthong & Brunfaut, 2020). Pedagogically, teachers can retain the coursebook's varied non-integrated tasks while supplementing them with integrated activities, such as listen-to-speak or listen-to-write tasks, to better simulate real-world communication.

Table 5. Task Types and Examples of *Listen This Way* (Level 4)

Integrated/ Non-integrated	Task type	Example(s)
Non-integrated	Dictation	**B** **Spot dictation. Fill in the blanks with the words you hear.** We have just two more days to "shop till you _____" till Christmas, and that typically means a lot of last minute _____. Online _____ is more popular than ever. Christmas shopping in the _____ is great sometimes. But when you've just got to _____ a bunch of shopping _____ and you want to have something to send to a _____ location, it's much better _____. In the U.S., online shopping is predicted to be _____ ahead of last December, while _____ can expect less than a 6% increase in sales. In _____, Europe's biggest e-tailing market, online _____ is expected to grow by _____, with $8 billion worth of goods sold through the _____ in the November to January _____.
Non-integrated	Blank-filling	**C** **Now listen to some tips for foreigners on how to do online-shopping in China. Listen carefully and supply the missing information to complete the summary.** **Summary:** China has a number of its own mature _____. If expats want to shop online in China, they can get _____ first, find the goods they want, and tell the websites their _____ and _____. You can pay with _____, through _____ or on _____. You can also ask for a refund if you find the products have _____ problems, although _____ differ from one website to another. And please note different websites have different _____. It is better to ask friends about their shopping experience before buying something from a specific website.

<table>
<tr><th>Integrated/ Non-integrated</th><th>Task type</th><th>Example(s)</th></tr>
<tr><td></td><td></td><td>
A Listen to the report. Supply the missing information.
National Teach Children to Save Day
Subjects: elementary school children across the country
Time: _____
Sponsor: American Bankers Association Education Foundation
Purpose: _____
Way of teaching: _____ bankers making _____

Dialogue two: Complete the chart. Use ticks to indicate whether the expenditures on those items are up, down or the same as last year. Supply the figures mentioned.
<table>
<tr><th>Items</th><th>Same</th><th>Up</th><th>Down</th><th>Figures</th></tr>
<tr><td>Entertainment</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Depreciation</td><td></td><td></td><td></td><td>300</td></tr>
<tr><td>Secretarial expenses</td><td></td><td></td><td></td><td></td></tr>
<tr><td>...</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Motor expenses</td><td></td><td>✓</td><td></td><td></td></tr>
<tr><td>Audit & accountancy costs</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Telephone & postage</td><td></td><td></td><td></td><td>Half as much as last year</td></tr>
</table>
</td></tr>
<tr><td>Non-integrated</td><td>Open-ended questions</td><td>
A The following passage is about a new way of payment. While listening for the first time, don't forget to take down some key words in the notes column. After the second listening, answer the following questions with key words.
Notes
_____ _____
_____ _____
_____ _____
_____ _____
_____ _____
Questions:
1. How can one pay for his groceries?
_____
2. What other functions does the finger have?
_____
3. How much of a finger does the store Thriftway use for payment?
_____
</td></tr>
</table>

Integrated/ Non-integrated	Task type	Example(s)
		In this section you will learn something about the personal savings situation in the United States. After listening, answer the following questions. 1. What is the title of the book the man is trying to promote? _____. 2. Among all the interviewees, how much is the biggest amount of money in their savings account? _____. 3. What are the three types of people in U.S.? a. _____; b. _____; c. _____. 4. According to Alan Greenspan, what is the ultimate goal of the Federal Reserve? _____. 5. According to the reporter, what will too much easy credit for too long result in? _____.
Non-integrated	True/False	Watch the video clip taken from an American romantic comedy film *Confessions of a Shopaholic* and decide whether the following statements are True or False. 1. _____ Little Rebecca envied grown-up girls who could get what they wanted with magic cards. 2. _____ When Rebecca grew into a big girl, she owned 9 bank cards. 3. _____ Someone had stolen Rebecca's credit card and bought something in Outdoor World. 4. _____ Rebecca owed banks sixteen thousand twelve hundred and sixty-two dollars and seventy cents. 5. _____ Suze thought it ironic for Rebecca to take a job at a savings magazine. 6. _____ Finally, Rebecca successfully controlled the urge to shop.
Non-integrated	Matching	**B** **Now listen again. The following are some of the tips for parents to foster the savings habit in their children. Match Column A (Purpose) with Column B (How to do).** **Column A** / **Column B** 1. Making savings visible and real — __ giving children play money to "invest" in stocks they can track in local newspapers 2. Encouraging children to save as much as they can — __ giving children an allowance and part of it going into their own savings 3. A first step toward learning to budget — __ putting 25 cents away for every dollar the children earn 4. Making savings and investing fun — __ building up savings in a piggy bank/opening children's own bank savings account

Integrated/ Non-integrated	Task type	Example(s)
Non-integrated	MCQs	**A** **Listen to a news report about safe investing and answer the questions.** 1. What is the first guideline to smart investing? A. Developing a savings plan. B. Setting up a bank account. C. Setting clear investing goals. 2. What is a "diversified portfolio" of investments according to Mr. Boros? A. A wide selection of investments. B. A limited range of growth stocks. C. A group of low risk bonds and securities. 3. Which kind of investment was NOT mentioned by Mr. Boros? A. Mutual funds. B. Bonds. C. Certificate of deposits.

Note. The task examples in this table are from *Listen This Way* (Level 4) (Higher Education Press, China).

Principle 6: Listening tasks should engage learners in both bottom-up and top-down processing

Based on Principle 6 and its indicators, the listening tasks in the four units of the coursebook were analyzed for the cognitive processes they engage. Tasks targeting bottom-up processing were identified, including dictation, which requires students to transcribe the spoken input verbatim, and blank-filling, which requires them to identify specific words and phrases (see Table 5). Tasks targeting top-down processing were also present, such as open-ended questions introduced by "why" and "how" prompts, which require students to synthesize information or draw inferences. Overall, the findings answer the guiding question affirmatively: the listening tasks engage learners in both bottom-up and top-down processing.

This pattern of task types is consistent with the view in L2 listening pedagogy that effective instruction should develop both bottom-up and top-down processing (Goh & Vandergrift, 2022; Vandergrift & Tafaghodtari, 2010). Pedagogically, teachers can make the targeted processing explicit to learners during instruction, thereby fostering learners' awareness of how bottom-up and top-down processing operate during listening and further supporting the development of overall listening ability.

Principle 7: Listening tasks should facilitate the development of metacognitive awareness

The listening tasks in the four units of the coursebook were analyzed for their coverage of metacognitive awareness across the pre-, while-, and post-listening phases, in accordance with Principle 7 and its indicators. It was found that each of the four units included a preparatory task introducing key vocabulary before the main listening activities and concluded with a review task consolidating the retention of important words and phrases (see examples in Table 6). While both tasks are pedagogically useful in engaging learners with pre-listening preparation and post-

listening review, they only focus on linguistic knowledge rather than the full dimensions of metacognitive awareness. The coursebook appears to lack tasks that prompt reflection on learners' strengths, weaknesses, or affective responses (person knowledge), preparation for the characteristics and demands of listening tasks (task knowledge), and the introduction and guided application of listening strategies (strategy knowledge and strategy use). These observations, therefore, offer a negative answer to the guiding question: the listening tasks do not substantively support learners in planning, monitoring, and evaluating their listening processes.

The limited attention to metacognitive awareness is consistent with the broader tendency in L2 listening instruction to prioritize comprehension outcomes over the processes through which comprehension is achieved (Goh, 2008; Goh & Vandergrift, 2022). Greater attention to metacognitive instruction is warranted, given the empirical evidence that such instruction improves L2 listening proficiency (Bozorgian & Shamsi, 2025; Vandergrift & Tafaghodtari, 2010) and the meta-analytic finding that metacognitive awareness is significantly correlated with listening ability (In'nami et al., 2022). Pedagogically, teachers can supplement the coursebook with activities such as listening diaries, guided pre- and post-listening reflection, and self-assessment checklists to foster learners' metacognitive awareness systematically.

Taken together, the illustrative example shows that the listening texts of *Listen This Way* (Level 4), particularly in the four units analyzed, are largely well-developed: they reflect features of authentic spoken language (Principle 1), include some multimodal input (Principle 3), and are broadly appropriate in difficulty level (Principle 4), although accent variety remains limited to British and American English (Principle 2). The listening tasks, however, are relatively limited: while tasks targeting both bottom-up and top-down processing are present (Principle 6), they are confined to non-integrated tasks (Principle 5), and the full dimensions of metacognitive awareness are largely unaddressed (Principle 7). The illustrative example demonstrates how the proposed framework supports a structured and systematic analysis of an L2 listening coursebook in ways that a generic, multi-skill framework would be less equipped to achieve.

Table 6. Examples of Preparatory and Review Tasks in *Listen This Way* (Level 4)

Task Type	Example
Preparatory task	**A** **The following words will appear in this unit. Listen carefully and study the definitions.** 1. ***e-tailing:*** the selling of goods and services through the Internet, also known as electronic retailing 2. ***register:*** to put one's name on an official list in order to be able to do something or to receive a service 3. ***refund:*** an amount of money that is given back to someone who has returned a product, paid too much, etc. 4. ***specialty:*** something that a person, a place or a company is known for making, producing or selling very well 5. ***cyberspace:*** the online world of computer networks and especially the Internet 6. ***dispatch:*** to send (someone or something) quickly to a particular place or for a particular purpose 7. ***biometrics:*** the measurement and analysis of unique physical or behavioral characteristics (as fingerprint or voice patterns) especially as a means of verifying personal identity
Review task	**Part VI Reminder of key points in this unit** Part I: register; e-tailing, specialty, refund policy, language barrier; ahead of Part II: e-mail, dispatch; obstacle, cyberspace, retail channel; Part III: verify, fingerprint; debit card, affiliate, biometrics, scramble, swirl, algorithm; nothing short of …, down the line, at one's service Part VI: get exercise; record album, rides, palm tree, neon, outlet mall, shoppertainment; athletic Part V: make profits, lure, line up; Black Friday, Christmas shopping season, the Thanksgiving holiday, discount, retail store, indication, electronics, flat screen TVs, economist; annual

Note. The task examples in this table are from *Listen This Way* (Level 4) (Higher Education Press, China).

CONCLUSION

Drawing on theories of language assessment, listening comprehension, and second language acquisition, this article proposes an evaluation framework for L2 listening coursebooks. The framework comprises seven principles, each addressing a specific dimension of listening texts or tasks. For each principle, a guiding question is provided to raise teachers' critical awareness of the relevant dimension, accompanied by a set of practical indicators to support its application. To illustrate how the framework operates in practice, we demonstrated its application to a specific L2 listening coursebook, *Listen This Way* (Level 4).

Designed specifically for L2 listening coursebook evaluation, the proposed framework has meaningful implications in practice. For teachers, this framework serves as a theory-based and accessible tool for evaluating and selecting L2 listening instructional materials, and fosters teachers' critical awareness of how such materials can be adopted and adapted to facilitate learners' L2 listening competence. For coursebook authors and publishers, the framework provides concrete guidance on the design of listening texts and tasks conducive to effective L2 listening instruction. It is also worth noting that generative AI holds great potential for designing and developing listening materials (Goh & Aryadoust, 2025). For example, large language models can assist in producing listening texts that meet specified requirements and in generating tasks that target different aspects of listening comprehension; speech synthesis can convert listening texts into speech with a range of English accents; and text-to-video tools can create multimodal listening materials that integrate spoken language with visual information. These capabilities could make it more time-efficient and cost-effective to develop listening texts and tasks that align with the principles proposed in this framework. Importantly, human oversight should always remain essential throughout the process.

Several limitations should be acknowledged. First, the framework is conceptual and exploratory in nature. Empirical validation remains necessary, for example, by examining how its application relates to learning and teaching outcomes. Second, this framework is intended for pre-instructional, diagnostic evaluation of L2 listening instructional materials and cannot replace the ongoing, context-based evaluation that effective teaching requires. As learner needs vary across contexts and evolve over time, teachers are encouraged to revisit their evaluation of instructional materials regularly. We hope this framework will provide a productive foundation for subsequent research on L2 listening coursebook evaluation and contribute to the broader development of language-teaching materials evaluation.

FUNDING

This study was supported by the Social Science Research Project of Chongqing Municipal Education Commission (grant number 24SKGH182) and Nanyang Technological University (NTU) Provost Graduate Award scholarship.

REFERENCES

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732935

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). Freeman.

Aryadoust, V. (2024). Topic and accent coverage in a commercialized L2 listening test: Implications for test-takers' identity. *Applied Linguistics*, *45*(5), 765–785. https://doi.org/10.1093/applin/amad062

Aryadoust, V., & Luo, L. (2023). The typology of second language listening constructs: A systematic review. *Language Testing*, *40*(2), 375–409. https://doi.org/10.1177/02655322221126604

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

Batty, A. O., & Suvorov, R. (2024). Visual cues and listening. In E. Wagner, A. O. Batty, & E. Galaczi (Eds.), *The Routledge handbook of second language acquisition and listening* (pp. 307–318). Routledge. https://doi.org/10.4324/9781003219552-27

Bax, S. (2012). *Text inspector* [Online text analysis tool]. https://textinspector.com/

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Bozorgian, H., & Shamsi, E. (2025). A review of research on metacognitive instruction for listening development. *International Journal of Listening*, *39*(1), 1–16. https://doi.org/10.1080/10904018.2023.2197008

Brown, J. D., & Crowther, D. (2023). *Shaping learners' pronunciation: Teaching the connected speech of North American English*. Routledge. https://doi.org/10.4324/9781003144779

Buck, G. (2001). *Assessing listening*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732959

Cubillos, J. H., Chieffo, L., & Fan, C. (2008). The impact of short-term study abroad programs on L2 listening comprehension skills. *Foreign Language Annals*, *41*(1), 157–186. https://doi.org/10.1111/j.1944-9720.2008.tb03284.x

Cunningsworth, A. (1995). *Choosing your coursebook*. Heinemann.

Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT Journal, 51*(1), 36–42. https://doi.org/10.1093/elt/51.1.36

Field, J. (2004). An insight into listeners' problems: Too much bottom-up or too much top-down? *System, 32*(3), 363–377. https://doi.org/10.1016/j.system.2004.05.002

Field, J. (2008). *Listening in the language classroom*. Cambridge University 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.

Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly*, *15*(3), 219–236. https://doi.org/10.1080/15434303.2018.1453816

Galloway, N., & Rose, H. (2018). Incorporating Global Englishes into the ELT classroom. *ELT Journal, 72*(1), 3–14. https://doi.org/10.1093/elt/ccx010

Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal, 58*(4), 363–374. https://doi.org/10.1093/elt/58.4.363

Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal, 39*(2), 188–213. https://doi.org/10.1177/0033688208092184

Goh, C. C. M., & Aryadoust, V. (2025). Developing and assessing second language listening and speaking: Does AI make it better? *Annual Review of Applied Linguistics*, *45*, 179–199. https://doi.org/10.1017/S0267190525100111

Goh, C. C. M., & Vandergrift, L. (2022). *Teaching and learning second language listening: Metacognition in action* (2nd ed.). Routledge. https://doi.org/10.4324/9780429287749

Harding, L. (2012). Accent, listening assessment and the potential for a shared-L1 advantage: A DIF perspective. *Language Testing*, *29*(2), 163–180. https://doi.org/10.1177/0265532211421161

Hill, D. A., & Tomlinson, B. (2013). Coursebook listening activities. In B. Tomlinson (Ed.), *Developing materials for language teaching* (2nd ed., pp. 429–442). Bloomsbury.

Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal, 48*(4), 315–328. https://doi.org/10.1093/elt/48.4.315

In'nami, Y., Koizumi, R., Jeon, E. H., & Arai, Y. (2022). L2 listening and its correlates: A meta-analysis. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigations* (pp. 235–283). John Benjamins. https://doi.org/10.1075/bpa.13.08inn

Jones, C. (2022). Authenticity in language teaching materials. In J. Norton & H. Buchanan (Eds.), *The Routledge handbook of materials development for language teaching* (pp. 65–77). Routledge. https://doi.org/10.4324/b22783-7

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: English language in the outer circle. In R. Quirk & H. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.

Kang, O., Yan, X., Kostromitina, M., Thomson, R., & Isaacs, T. (2024). Fairness of using different English accents: The effect of shared L1s in listening tasks of the Duolingo English test. *Language Testing*, *41*(2), 263–289. https://doi.org/10.1177/02655322231179134

Karakaş, A. (2024). Supporting the teaching of global Englishes language teaching (GELT) through Voki. *RELC Journal*, *55*(3), 876–880. https://doi.org/10.1177/00336882231199515

Krashen, S. D. (1981). *Second language acquisition and second language learning* (1st ed.). Pergamon Press.

Kwon, S. K., & Yu, G. (2024). The effect of viewing visual cues in a listening comprehension test on second language learners' test-taking process and performance: An eye-tracking study. *Language Testing*, *41*(3), 649–680. https://doi.org/10.1177/02655322241239356

Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed., pp. 179–211). Cambridge University Press. https://doi.org/10.1017/9781139042789.011

Matsuda, A. (2019). World Englishes in English language teaching: Kachru's six fallacies and the TEIL paradigm. *World Englishes*, *38*(1–2), 144–154. https://doi.org/10.1111/weng.12368

McCarthy, M. (2010). Spoken fluency revisited. *English Profile Journal*, *1*, 1–15. https://doi.org/10.1017/S2041536210000012

McCarthy, M., & Carter, R. (1995). Spoken grammar: What is it and how can we teach it? *ELT Journal*, *49*(3), 207–218. https://doi.org/10.1093/elt/49.3.207

McDonough, J., Shaw, C., & Masuhara, H. (2013). *Materials and methods in ELT: A teacher's guide* (3rd ed.). Wiley-Blackwell.

Mukundan, J., & Ahour, T. (2011). A review of textbook evaluation checklists across four decades (1970–2008). In B. Tomlinson & H. Masuhara (Eds.), *Research for materials development in language learning: Evidence for best practice* (pp. 336–352). Bloomsbury Publishing Plc. https://doi.org/10.5040/9781474211949.ch-021

Ockey, G. J., & Wagner, E. (Eds.). (2018). *Assessing L2 listening: Moving towards authenticity*. John Benjamins. https://doi.org/10.1075/lllt.50

Renandya, W. A., & Hu, G. (2018). L2 listening in China: An examination of current practice. In A. Burns & J. Siegel (Eds.), *International perspectives on teaching the four skills in ELT* (pp. 37–50). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-63444-9_3

Richards, J. C. (2005). Second thoughts on teaching listening. *RELC Journal*, *36*(1), 85–92. https://doi.org/10.1177/0033688205053484

Richards, J. C. (2014). The ELT textbook. In S. Garton & K. Graves (Eds.), *International perspectives on materials in ELT* (pp. 19–36). Palgrave Macmillan. https://doi.org/10.1057/9781137023315_2

Rukthong, A. (2021). MC listening questions vs. integrated listening-to-summarize tasks: What listening abilities do they assess? *System*, *97*, 102439. https://doi.org/10.1016/j.system.2020.102439

Rukthong, A., & Brunfaut, T. (2020). Is anybody listening? The nature of second language listening in integrated listening-to-summarize tasks. *Language Testing*, *37*(1), 31–53. https://doi.org/10.1177/0265532219871470

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. https://doi.org/10.1111/j.0023-8333.2005.00320.x

Suvorov, R. (2015). The use of eye tracking in research on video-based second language (L2) listening assessment: A comparison of context videos and content videos. *Language Testing*, *32*(4), 463–483. https://doi.org/10.1177/0265532214562099

Syrbe, M., & Rose, H. (2018). An evaluation of the global orientation of English textbooks in Germany. *Innovation in Language Learning and Teaching, 12*(2), 152–163. https://doi.org/10.1080/17501229.2015.1120736

Taylor, L., & Geranpayeh, A. (2011). Assessing listening for academic purposes: Defining and operationalising the test construct. *Journal of English for Academic Purposes, 10*(2), 89–101. https://doi.org/10.1016/j.jeap.2011.03.002

Tomlinson, B. (Ed.). (2013). *Developing materials for language teaching* (2nd ed.). Bloomsbury.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, *60*(2), 470–497. https://doi.org/10.1111/j.1467-9922.2009.00559.x

Vettorel, P., & Lopriore, L. (2013). Is there ELF in ELT coursebooks? *Studies in Second Language Learning and Teaching, 3*(4), 483–504. https://doi.org/10.14746/ssllt.2013.3.4.3

Wagner, E. (2008). Video listening tests: What are they measuring? *Language Assessment Quarterly*, *5*(3), 218–243. https://doi.org/10.1080/15434300802213015

Wagner, E. (2010). Test-takers' interaction with an L2 video listening test. *System*, *38*(2), 280–291. https://doi.org/10.1016/j.system.2010.01.003

Wagner, E. (2014). Using unscripted spoken texts in the teaching of second language listening. *TESOL Journal*, *5*(2), 288–311. https://doi.org/10.1002/tesj.120

Wagner, E. (2018). Texts for listening instruction and assessment. In J. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1544–1555). Wiley-Blackwell. https://doi.org/10.1002/9781118784235.eelt0626

Wagner, E., & Toth, P. D. (2014). Teaching and testing L2 Spanish listening using scripted vs. unscripted texts. *Foreign Language Annals*, *47*(3), 404–422. https://doi.org/10.1111/flan.12091

Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan. https://doi.org/10.1057/9780230514577

Yeldham, M., & Gruba, P. (2014). Toward an instructional approach to developing interactive second language listening. *Language Teaching Research*, *18*(1), 33–53. https://doi.org/10.1177/1362168813505395