ISSN 2078-5119. Теорія і практика викладання української мови як іноземної. 2026. Випуск 20. С. 131–145
Theory and Practice of Teaching Ukrainian as a Foreign Language. 2026. Issue 20. P. 131–145

ІІ. ФОРМУВАННЯ МОВНОЇ ТА МОВЛЕННЄВОЇ КОМПЕТЕНЦІЇ В ПРОЦЕСІ ВИВЧЕННЯ УКРАЇНСЬКОЇ МОВИ ЯК ІНОЗЕМНОЇ

УДК 371.26:811.161.2'243:37.016

TOWARDS THE IDEAL LISTENING INSTRUCTION: CROSS-TEST INSIGHTS FROM HIGH-STAKES ASSESSMENTS

Kseniia Borodin

Ukrainian Catholic University
School of Ukrainian Language and Culture
17 Sventsitskoho St., 79000, Lviv, Ukraine
phone: 097 822 66 75
email: borodinka@ucu.edu.ua
https://orcid.org/0009-0001-9307-1363

The article analyzes the features of task instruction formulation in the listening subtest of high-stakes language examinations, including the English-language tests IELTS and TOEFL, the German TestDaF, the Czech Certifikovaná zkouška z češtiny pro cizince (CCE), and the Polish national exam Państwowy Egzamin Certyfikatowy z Języka Polskiego jako Obcego. A comparative analysis demonstrated that the modes of presenting instructions in the international tests differ significantly in structure, length, and level of communicative density. Several principal models were identified – ranging from highly concise procedural instructions to more extended communicatively oriented forms designed to create an authentic context of interaction.

Special attention is given to comparing the content of task instructions with the expectations outlined in the Common European Framework of Reference for Languages (CEFR). The comparison revealed that instructions in most tests are primarily shaped by task type rather than by the targeted language proficiency level. As a result, they often fail to fully reflect the expected cognitive complexity or the types of speech acts characteristic of corresponding CEFR levels.

To examine the impact of linguistic formulation on task comprehension, an empirical study was conducted using Ukrainian-language test instructions. The experiment analyzed participants' reactions and responses, which helped to identify typical difficulties related to cognitive load, excessive complexity, or, conversely, oversimplification that can reduce the authenticity of the communicative situation.

The findings confirm that the clarity, structural organization, and linguistic transparency of instructions directly influence performance in listening tasks. Based on the analysis, a set of practical recommendations has been formulated for developers of Ukrainian as a foreign language tests. These include principles for drafting instructions that ensure a balance between accessibility, cognitive optimality, and communicative relevance. The proposed recommendations aim to enhance contextual validity, reduce the impact of construct-irrelevant factors on assessment results, and improve the overall reliability of language testing.

*Key words***:** listening comprehension, task instructions, Ukrainian as a foreign language, assessment, contextual validity.

DOI: http://dx.doi.org/10.30970/ufl.2026.20.5117

Introduction. A frequently asked question in educational forums is: *"Which international English language test is better – IELTS or TOEFL?"* On one Ukrainian informational platform, the answer was formulated as follows: *"TOEFL is considered more predictable, while IELTS uses a wider variety of task types, especially since the instructions for tasks are constantly different"* (edusteps.com.ua).

This comment drew attention to a relatively understudied yet important aspect of language assessment – the role of task instructions, which can significantly influence test-taker orientation and performance. Although often regarded as secondary, instructions

form an essential component of the test construct and deserve systematic investigation.

The effectiveness of task performance in a language test depends not only on the candidate's level of language proficiency, but also significantly on how clearly they understand the task instructions. In any form of testing, task instructions play a critical role in ensuring construct validity, procedural fairness, and linguistic accessibility. Unlike informal classroom activities, high-stakes exams – such as the English-language TOEFL and IELTS, the German TestDaF, the Czech CCE, or other national certification tests – carry significant consequences for the candidate's academic, professional, or migratory future. Well-formulated instructions serve not only an explanatory, but also a motivational function: they orient the candidate toward the purpose of the task, reduce cognitive load, and enable a greater focus on the communicative activity itself. As such, the clarity, structure, and function of task instructions directly influence task performance outcomes [8 : 98; 4 : 122].

Listening tasks in particular pose specific cognitive challenges. Listening occurs in real time, without the opportunity to review or clarify information. This is why the development of clear, concise, and communicatively effective instructions is a key component of high-quality test design.

Formulation of the problem. While it may be assumed that instructions for listening tasks in high-stakes language exams are largely standardized, with only minor national variations, in practice they demonstrate considerable variability depending on the conceptual model adopted by each test. Consequently, there is a clear need to analyze existing approaches – their advantages and limitations – in order to justify an optimal model for instruction design in the context of Ukrainian language testing.

The aim of this article is to develop a model of the ideal instruction for listening tasks based on a comparative analysis of instruction formats used in leading high-stakes language exams, such as IELTS, TOEFL, and TestDaF, as well as in national certification tests in Poland (Państwowy Egzamin Certyfikatowy z Języka Polskiego jako Obcego) and the Czech Republic (Certifikovaná zkouška z češtiny pro cizince – CCE). By analyzing the structural, linguistic, and functional characteristics of task instructions in these exams, the study seeks to offer recommendations for creating instructions for a Ukrainian language test that are both communicatively authentic and cognitively accessible. The broader goal is to support the development of instruction designs that enhance test validity, reduce anxiety levels, and ensure equitable conditions for a diverse population of test-takers.

Analysis of recent research and publications. Scholarly literature commonly identifies at least three core functions of test instructions: organizational, communicative, and metacognitive. The most fundamental function of instructions lies in structuring the test – outlining the task format, expected response type, time constraints, number of listenings, and so forth. As Bachman and Palmer note, clearly structured instructions reduce extraneous cognitive load and allow test-takers to focus more effectively on language processing [2 : 181–188].

Instructions also can serve a communicative role by providing context: they describe the situation, define the roles of the speakers, and indicate the genre of the text. This facilitates the activation of relevant schemata and the formation of pragmatic expectations, which are especially critical for tasks that simulate real-life communicative settings [5 : 40]. This approach enhances the ecological validity of a test.

Moreover, instructions can perform a metacognitive function by guiding the test-taker toward effective strategies: what to focus on, when to shift attention, and how to anticipate

key information. Vandergrift and Goh emphasize that successful listening comprehension depends not only on linguistic skills but also on metacognitive regulation [9 : 13]. Well-designed instructions support this by embedding cues for goal-setting and self-monitoring.

Thus, achieving a balance between formality and contextualization emerges as a key challenge in the design of test instructions.

Main body. This study employs a comparative qualitative analysis of listening task instructions in five high-stakes language proficiency tests: IELTS, TOEFL, TestDaF, as well as two national certification exams in Poland and the Czech Republic. The aim is to identify various models and features that contribute to the communicative effectiveness, clarity, and accessibility of task instructions, and to develop a model of the "ideal" instruction for listening tasks.

The data corpus includes both oral and written instructions presented to candidates during the listening component. Instructions were collected from publicly available sources – namely, sample test materials, official guidelines, and published audio transcripts provided by test developers. As such, only instructions that were accessible in open-access resources were included in the analysis; materials not publicly disclosed were excluded.

1.1. Introductory instruction

In certification language exams, the listening subtest is accompanied by a multi-layered system of instructions, which vary by delivery mode (written vs. oral) and by timing within the subtest (introductory, transitional, and closing). These instructions may be either duplicated across modes or differ in content and emphasis.

The introductory instruction aims to familiarize candidates with the main parameters and organizational conditions of the subtest. It typically appears in two forms: a written instruction, printed on the cover page of the test booklet, and an oral instruction, delivered at the beginning of the listening component. Common elements of the written portion include:

- The name of the exam, the proficiency level, and the subtest title.
- Administrative guidelines: *"Write your name and candidate number in the spaces at the top of this page. Listen to the instructions for each part of the paper carefully"*; *"Answer all the questions on the basis of what is stated or implied by the speakers in this test"*; *"While you are listening, write your answers on the question paper"*; *"You will have 10 minutes at the end of the test to copy your answers onto the separate answer sheet."*
- Behavioral rules: *"Do not open this question paper until you are told to do so. Use a pencil"* (TOEFL); *"Wörterbücher oder Mobiltelefone nicht erlaubt"*; *"Během testu nesmíte mluvit s ostatními kandidáty a nesmíte používat slovník."*
- Timing information: *"Prüfungszeit für nicht behinderte Kandidatinnen und Kandidaten: 40 Minuten…"*

These written instructions are typically formalized and serve to inform candidates about the test structure, procedural rules, and restrictions.

The **oral instruction,** by contrast, contains components such as:

- Official identification: name of the exam body, the exam title, level, and subtest name (e.g., *"Państwowa Komisja do spraw poświadczania znajomości języka polskiego jako obcego. Egzamen certyfikatowy z języka polskiego jako obcego. Poziom B1. Rozumienie ze słuchu."*)
- Purpose of the subtest: often framed positively to reduce anxiety and highlight the candidate's opportunity to demonstrate listening skills (e.g., *"In the Listening Comprehension section of the test, you will have an opportunity to demonstrate your ability to understand spoken English."* – TOEFL)

• Structure overview: number of parts, number of questions, how many times each part is played, and total duration (e.g., *"There are four parts to the test. You will hear each part once only. There are 40 questions."* – IELTS)

• Procedural guidance: general: *"Proszę uważnie słuchać nagrań"*, *"Lesen Sie jeweils zuerst die Aufgaben und hören Sie dann den Text dazu."*; strategic use of pauses: *"For each part of the test, there will be time for you to look through the questions and time for you to check your answers."* (TOEFL)

• Scoring: *"Each question carries one mark"*; *"Každý úkol má pouze jedno správné řešení."*

• Technical aspects: audio signal before each recording: *"Przed każdym nagraniem usłyszą Państwo sygnał."*; pre-test audio check: *"Zkontrolujte, zda je zvuk v pořádku."*

• Initial overview time: *"Uwaga! Proszę teraz obejrzeć cały test."*

• Official test launch: *"Uwaga! Zaczynamy test."*

Across all tests, oral instructions tend not to replicate the written ones verbatim. The only universally consistent elements are the exam title, subtest name, and level; all other content is left to the discretion of the testing authority.

Thus, the introductory instruction in listening certification tests performs a vital role: it prepares candidates for the subtest by conveying the structure, expectations, and technical parameters – fostering a fair, organized, and procedurally transparent testing environment. Both written and oral components have distinct functional purposes: they inform about procedures, test structure, time limits, scoring criteria, permitted materials, and behavioral norms. Oral instruction does not merely duplicate the written text—it highlights critical points to ensure candidates begin on equal footing. This helps reduce anxiety and procedural confusion, ultimately supporting the validity of the test outcomes.

However, a full compilation of all aforementioned elements would result in an excessively long instruction – over 250 words – taking up to three minutes of audio time, or approximately one-tenth of a 30-minute subtest. Therefore, it is advisable to prioritize only the most essential content in oral delivery. Less critical information should be presented on the test booklet's cover, and the least critical in candidate information booklets or the exam website.

The following five components are proposed as mandatory for oral delivery in the introductory instruction:

1.Exam and subtest identification – this functions as the formal opening, orienting the candidate to the assessment context.

2.Subtest structure – number of parts and questions, number of times each part will be played, and total duration.

3.Procedural guidance – when to listen, when to write, and when to preview questions.

4.Answer sheet instructions – where and how to write answers, and whether transfer time will be provided.

5.Official start announcement – to clearly signal the beginning of the test.

All other content should be included only as necessary, based on the test format and delivery constraints, to avoid overloading the candidate.

In sum, a well-designed, concise introductory instruction sets the stage for valid and reliable testing by helping candidates focus, orient themselves, and mentally engage – even before the first audio prompt is heard.

1.2. Intermediate instructions

In modern listening tests, such as the IELTS and CCE, there is a clearly observable functional distinction between the oral intermediate instruction that the candidate hears and the written task wording presented in the test booklet. Both forms serve complementary functions, though their communicative focuses differ. The oral instruction is predominantly organizational and contextual – it introduces the candidate to the listening situation, outlines the genre and situational parameters of the text, and reminds them of the technical aspects of task performance. In contrast, the written instruction plays a normative role: it precisely formulates the action the test-taker is expected to perform.

For example, in **IELTS** a typical oral instruction sounds like: *"Listen to the conversation between a Japanese student and a housing officer and complete the form. First you have some time to look at questions 1 to 5."* This example introduces the context (a conversation between a student and a housing officer), specifies the type of task (completing a form), and allocates time for previewing the questions – an important part of strategic listening.

In subsequent sections, similar instructions might sound like: *"You will hear a talk given by a tour guide..."; "You will hear a discussion about shopping habits..."; "You will hear a lecture about study..."* Each time, the listener is guided by receiving cues about the situation, the text's register, and speaker roles. The accompanying written instruction, however, is much more formalized: *"Complete the notes. Write NO MORE THAN TWO WORDS OR A NUMBER for each answer"; "Choose the correct letter, A, B, or C"; "Complete the table. Write ONE WORD ONLY for each answer"; "Write the correct letter, A–G, next to the questions"*. Such wording clearly identifies the task type, the answer format, and sets limits on the form of the response (word count or use of numbers).

Thus, IELTS utilizes a dual-mode instruction system with complementary functions. Thanks to this standardization, instructions minimize the risk of misinterpretation and contribute to the objectivity of the assessment.

A similar approach can be observed in other international tests. In the German **TestDaF**, the instructions also have a clear functional structure, typically including: type of text (monologue, dialogue, interview), number of listenings (once or twice), type of task (e.g., multiple choice, matching), time for previewing the questions. A notable feature is the provision of context, which improves situational understanding.

Structurally, the instruction always consists of two parts: a very concrete formal task plus contextualization. For example: *"Sie hören nun eine Diskussion. Sie hören die Diskussion zweimal. Dazu lösen Sie acht Aufgaben. Ordnen Sie die Aussagen zu: Wer sagt was? Lesen Sie jetzt die Aussagen 23 bis 30. Dazu haben Sie 60 Sekunden Zeit. Der Moderator der Radiosendung 'Diskussion am Abend' diskutiert mit den Eltern Dana Schneider und Florian Bader zum Thema 'Sollen kleine Kinder in die Kinderkrippe gehen?"* This gives the candidate a full situational "map" before the audio starts, significantly reducing cognitive load.

The **Czech** format, similar to TestDaF, follows this two-component logic – beginning with a general directive followed by context (description of the situation and type of text). Example: "*Uslyšíte pět krátkých textů na různá témata. U každého úkolu 1–5 vyberete jednu správnou odpověď z nabídky A–C. Každou nahrávku uslyšíte dvakrát. Teď si přečtěte úkoly 1–5. Uslyšíte rozhovor s odborníkem na gastronomii."* Here too, the basic parameters are stated (number of texts, number of listenings, answer format), while the contextual part comes later and is more concise: *"Uslyšíte rozhovor s odborníkem na gastronomii."*

This shorter contextual framing, compared to the German format, may require more effort from the candidate during the initial orientation phase.

TOEFL, by contrast, tends toward minimal oral instructions. For instance, a candidate might hear: *"In Part A, you will hear short conversations between 2 people... After each conversation, you will hear a question..."* Such instruction focuses more on describing the format of the listening section rather than on providing a communicative context. It is procedural rather than contextually integrated, exemplifying a standardized model with a low level of contextualization.

In the **Polish test**, the structure of the instruction also comprises two parts. The first (formal) part is nearly always repeated across tasks and levels: *"Proszę uważnie słuchać tego nagrania i wykonywać zadanie zgodnie z podanym przykładem."* The second (contextualizing) part is not always present – especially in global listening tasks. When it is included, it typically mentions the genre: *"Proszę wysłuchać kilku krótkich tekstów"*. Or the topic: *"Proszę wysłuchać rozmowy na temat zabezpieczeń banknotów."* Note also the grammatical expectation, often stated explicitly: *"Forma gramatyczna odpowiedzi powinna być dopasowana do pytania."*

The analysis of intermediate instructions in listening tests demonstrates the effectiveness of combining two levels: contextual (oral) and formal (written). This approach helps lower candidates' anxiety, ensure correct comprehension of the listening situation, increase focus, and reduce the risk of procedural mistakes not related to language proficiency. In tests where context is not provided orally, it is advisable to include at least a written contextual description.

1.3. Final instructions

Final instructions are usually brief. They inform the test-taker about the completion of each task (e.g., *"Konec úlohy číslo čtyři"*) or the entire subtest, and if necessary, they remind candidates about the time available to check and transfer answers to the answer sheet. These instructions serve an important organizational and structuring function.

2. Lexico-grammatical analysis of instructions in Slavic-language tests

From the perspective of test fairness, instructions should be linguistically accessible to speakers of different languages: free from grammatical complexity, idiomatic expressions, or culturally specific references [5 : 146]. Research shows that unclear or overloaded instructions can create construct-irrelevant variance – that is, they affect results not due to linguistic proficiency, but due to comprehension difficulties [1 : 55]. In high-stakes international testing, even minor misunderstandings can lead to systemic distortions in test outcomes [7 : 49].

To develop a model for Ukrainian listening test instructions, only the Polish and Czech tests were analyzed – as they are grammatically and lexically closer to Ukrainian than English or German.

2.1 Grammatical Structure Analysis

This analysis reveals the **syntactic complexity** of instructions, including sentence length, use of active/passive voice, modal constructions (imperatives, conditionals), and overall grammatical clarity or density.

In the Polish test, instructions generally feature simple grammatical structures that

support ease of access and comprehension. Syntax is dominated by simple or compound sentences with clearly marked imperatives, such as: *"Proszę teraz obejrzeć…"*, *"Zaczynamy test"*, *"Proszę uważnie słuchać…"*. These are formed in the indicative or imperative mood and in active voice, avoiding passive constructions, which reduces cognitive load.

At higher levels, more complex structures appear, such as: *"Proszę wysłuchać tekstu i zdecydować, która z roślin (A–E) odpowiada podanym poniżej opisom."*

Despite the increased grammatical complexity, such instructions remain comprehensible due to familiar structure and topic-limited vocabulary.

Czech test instructions also use simple or complex sentences, e.g.: *"Uslyšíte rozhovory lidí, kteří volají na různá místa."* They frequently use imperative formulas with verbs like: *"Uslyšíte…", "Vyberete…", "Rozhodněte…"*, or indicative forms like *"Teď si přečtěte…"*. While syntactically a bit more complex than Polish, the Czech instructions remain well balanced.

Based on this grammatical analysis, the following practical recommendations for writing Ukrainian listening test instructions can be proposed to improve accessibility, clarity, and reduce cognitive load:

1.Use indicative mood for beginner levels (A1–A2). Avoid complex syntax for lower-level learners. Phrases like *"Зараз ви будете слухати діалог"* or *"Вам треба виконати завдання за зразком"* are easier to understand than passive or subordinate clauses.

2.Prefer imperative mood from B1 and above. Imperatives clearly signal the action to be taken by the candidate. Examples like *"Прослухайте текст", "Запишіть відповідь",* and *"Виберіть правильний варіант"* work well across levels starting from B1 and do not add unnecessary cognitive burden.

3.Limit the use of subordinate clauses. Instructions like *"Визначте, яка з рослин відповідає наведеному опису"* are only advisable if the vocabulary is simple and the topic is clear. Use them mostly from B2 level onward, and only when such syntax is more natural and doesn't hinder comprehension.

4.Syntactic simplicity as a guiding principle. Use simple or compound sentences with a direct word order. Avoid embedded clauses, double negation, complex conditionals, or passive constructions that may impair understanding.

5.Short and concise formulations. Each sentence should convey only one instruction or action. For example: *"Прослухайте приклад. Потім виконайте завдання 1–5."* This sequence is easier to follow and improves orientation.

2.2 Lexical Complexity

This assesses the lexical load based on word frequency, presence of specialized terms, and suitability for candidates at specific CEFR levels.

The vocabulary used in Polish and Czech test instructions is based almost exclusively on high-frequency words. Common, transparent words are used, such as: *nagranie, test, przykład, zadanie* – in Polish; *rozhovor, text, otázka, odpověď* – in Czech.

Answer formats such as *"z nabídky A–C"*, *"ANO / NE"*, or *"napište jedno správné slovo nebo číslo"* are transparent and familiar to candidates.

Importantly, in test instructions, semantic load is often concentrated in verbs, which define the actions the candidate must take. The most frequent verbs include: *прочитайте, прослухайте, виберіть, напишіть*. Less frequent but common ones include:

виконайте, позначте, поєднайте, вставте, відповідайте, перевірте, продовжіть. To better understand the function of these verbs, they can be interpreted within speech acts aligned with CEFR levels:

- **A1:** Instructions for recognition and repetition: *прочитайте, прослухайте, дайте відповідь.*
- **A2:** Add basic actions: *виберіть, поєднайте, напишіть.*
- **B1:** Include cognitive tasks: *заповніть, опишіть, поясніть.*
- **B2:** Introduce reasoning: *обґрунтуйте, порівняйте, проаналізуйте.*
- **C1:** Expect abstract operations: *інтерпретуйте, сформулюйте, оцініть, доведіть.*

At **A2** and **B1** levels in the Polish test, instructions are repeated with a limited set of verbs: *słuchać, zaznaczać, wybrać, wpisać, połączyć.* These verbs align with perceptual and motor actions without requiring complex mental operations. The common *"Proszę"* + *infinitive* structure softens the command and maintains clarity while reducing emotional stress.

In Czech, already at these levels, instructions use a broader range of verbs involving cognitive operations: *Rozhodněte, napište, vyberte, přečtěte, přiřaďte.*

The imperative is used without softening but remains formally structured. These instructions involve decision-making and matching – implying a higher cognitive demand than the Polish test at **A2**.

At **B2**, Polish instructions begin to include new verbs such as: *odpowiedzieć, zdecydować, dopasować.* These reflect reflective information processing – candidates must now formulate responses, make decisions, and justify choices.

In contrast, Czech instructions remain lexically stable, maintaining the same verbs across levels.

At **C1**, the diversity of verbs increases: in Polish – *połączyć, wpisać, dopasować, zdecydować* – indicating integration, analysis, and data transformation. In Czech: *doplňte* – indicating lexical production rather than just recognition or selection.

Despite this, the overall structure of instructions remains consistent with earlier levels – maintaining coherence and saving cognitive resources by avoiding new grammar structures.

However, analysis of Polish and Czech tests shows that verb choice is driven more by task type than CEFR level. For instance, a matching task at any level will likely use *přiřaďte* (Czech) or *dopasuj* (Polish). Verbs like *zaznaczyć, wybrać, napište, rozhodněte* are used across A2–C1 levels regardless of the language proficiency expected.

To examine how different formulations of task instructions influence test-takers' comprehension, confidence, and performance in high-stakes listening assessments, an exploratory mixed-methods study was conducted. A total of 30 participants (N = 30), representing a range of linguistic backgrounds and CEFR proficiency levels (A1 to C1), were asked to evaluate three versions of listening instructions. The three prompts varied in length, specificity, and cognitive load: These instructions varied in terms of length, specificity, and cognitive demand. Instruction 1 was brief and minimalist: *«Виберіть правильну відповідь до кожного питання. Ви почуєте запис двічі.»* Instruction 2 offered moderate detail with contextual grounding: *«Ви слухаєте розмову між студентом і викладачем. Виберіть правильну відповідь до кожного питання. Запис буде програно двічі.»* Instruction 3 was the most elaborate, explicitly stating timing, procedures, and discourse context: *«Ви почуєте діалог між студентом і викладачем щодо вибору*

курсу. Завдання: виберіть правильну відповідь до кожного з п'яти питань. У вас буде 30 секунд на ознайомлення. Запис пролунає двічі.»

Participants were asked to indicate their preferred version, explain the rationale for their choice, describe how they would approach the task after reading the instruction, and identify any unclear or missing information. Data were analyzed both quantitatively (frequency of preference) and qualitatively (thematic analysis of open responses).

The majority of participants expressed a clear preference for Instruction 2, which provided a balance between brevity and contextual grounding. It was perceived as sufficiently informative to prepare the listener for the task without overwhelming them with procedural detail. Phrases like "short, clear and easy to understand," "a good combination," and "the clearest" were frequent across responses, emphasizing the importance of cognitive accessibility.

In contrast, Instruction 1, although praised for its clarity and simplicity, was often seen as lacking critical context—particularly regarding the nature of the recording and the number of questions to expect. Instruction 3, while the most comprehensive, elicited negative feedback from many respondents, who reported increased anxiety, cognitive overload, and a sense of distraction due to its verbosity.

In several cases, participants noted that some background about the speakers or situation helped them focus or prepare cognitively, especially when the instruction included social context (e.g., "official style conversation," "student and teacher"). This aligns with established findings in listening comprehension research: contextual pre-activation supports better inferencing and strategic listening.

Notably, instruction preferences appeared to correlate with both language proficiency and cognitive style. Advanced speakers (C1) and participants demonstrating greater metacognitive awareness tended to favor more concise and minimal instructions (Instruction 1), as they required less scaffolding. In contrast, lower-proficiency users (A1–B1) expressed a greater need for clarity, structure, and orientation, showing a preference for more detailed formats (Instruction 2 or 3). Participants who self-identified as neurodivergent, or who emphasized a personal need for predictability and cognitive grounding, also favored context-rich instructions, particularly Instruction 3.

When asked whether their instruction preference would change for a more difficult or longer listening task, responses revealed two key trends. First, a substantial portion of participants (approximately half) stated that their preference would remain unchanged – often citing a continued desire for clarity, brevity, and emotional comfort. Statements such as "No, I like clear, concise instructions communicated with the fewest words possible" and "No, I think I would prefer this instruction even if the task was longer or more difficult" reflect a stable cognitive style that favors simplicity and minimal processing load regardless of task difficulty.

However, a significant minority (about 35–40%) indicated that their preference might shift in response to increased task demands. These participants leaned toward preferring Instruction 3 in more challenging or extended tasks, arguing that more procedural and contextual detail would help them manage complexity. For instance, one respondent noted that "if it were longer than I might choose number 1 because it seems easier," while another said, "perhaps [Instruction] 3 is better for a long dialogue because I know there will be only five questions." These answers suggest that when facing a longer or cognitively demanding input, some test-takers value predictive scaffolding, such as knowing the number of speakers or the amount of content to anticipate.

Across the sample, several informational elements consistently emerged as critical to test-takers: confirmation that the recording would be played twice, clarity about the amount of preview time available before the listening task, and explicit identification of the communicative context (e.g., student–teacher dialogue, airport announcement).

Despite overall comprehension of the task, participant responses revealed varying levels of strategic readiness for the listening phase. Some respondents demonstrated effective use of the 30-second preview window, such as by skimming questions, highlighting keywords, or anticipating likely content based on context (e.g., "announcement," "at the airport"). However, many participants exhibited limited or passive strategies – for example, simply "waiting for the audio," "listening and hoping to catch the answers," or experiencing stress due to time pressure. Several responses indicated uncertainty about whether the 30 seconds were meant for reading, listening, or answering, suggesting that even well-written instructions may fail to activate efficient listening behavior unless learners are explicitly trained in how to use preparatory time.

This gap between understanding the instruction and employing optimal listening strategies underscores the importance of metacognitive scaffolding in test preparation. Simply knowing the recording will be played twice, or that the task involves a dialogue, does not necessarily translate into intentional cognitive preparation—especially for lower-proficiency or anxious test-takers.

One of the most frequently discussed elements in participant responses was the explicit reference to the 30-second preview period, which Rita Green considers an obligatory component [5 : 89]. For many respondents, this detail was perceived as *reassuring and informative* – helping them understand the pacing of the task, allocate attention, and prepare strategically. Particularly among higher-proficiency (B2–C1) users and those with metacognitive awareness, knowing that there was time for previewing was seen as a helpful cognitive scaffold.

However, a notable subset of participants – especially among lower-proficiency users (A1–B1) and those reporting heightened stress responses – described the mention of the 30-second window as a source of anxiety or pressure. These users interpreted the time constraint not as an opportunity, but as a challenge or demand. In some cases, participants were even unsure when exactly the 30 seconds applied (before, after, or during the listening), revealing that the same piece of information may enhance clarity for some but raise uncertainty or stress for others.

These findings support the hypothesis that optimal listening instructions must be neither too sparse nor excessively detailed, but rather tailored to cognitive accessibility and user expectation. The study also suggests a potential benefit in offering differentiated instruction formats or adaptive guidance, depending on the test-taker's profile, linguistic background, and anxiety sensitivity. Such considerations are particularly critical in high-stakes testing environments, where instructional clarity directly impacts performance and fairness.

To accommodate both profiles, instructions should retain the time information but frame it in a more supportive and neutral way. For example: *"You will have a short period (30 seconds) to look over the questions before the recording begins. Use this time to understand what to listen for."* Such wording communicates the key information without emphasizing time pressure. This approach offers structure for those who need it, while reducing the threat for more anxious test-takers – and is aligned with universal design principles in assessment.

A particularly insightful comment highlighted a strategic shift in listening behavior depending on audio length: *"With short audios, I scramble to hear key words and answer questions while listening. With a longer audio, I'm more likely to listen and try to understand the conversation so that I can recall it and answer the questions after."* This reveals a metacognitive awareness of how task length influences listening strategies, and suggests that more structured instructions might support a deeper processing mode in extended tasks.

Finally, one respondent noted that instruction 3 "would become more firm" in a high-stakes or long-task context, implying that test-takers may tolerate or even prefer denser instructions when the stakes or task complexity justify them.

In sum, while concise and contextually balanced instructions (like Instruction 2) remain broadly preferred, task difficulty and length can shift user needs toward more detailed formats. This reinforces the need for adaptive or differentiated instruction design, particularly in high-stakes listening assessments where cognitive load and test anxiety interact dynamically with user preferences.

When participants were asked what they would do after reading the listening task instructions, their responses highlighted considerable variation in strategic behavior. Some respondents demonstrated a clear understanding of how to prepare for a listening task, indicating they would:

- Use the preview time effectively to read questions and identify keywords: *"I would read the questions that went with the recording and prepare myself to listen for that information"; "Я буду підкреслювати такі ключові слова, наприклад, місце в аеропорту"; "I would read the questions during the 30 second period that I have been given to do so."*
- Visualize or anticipate the context: "*Think about some situations that could happen in an airport"; "Read the sentence, so I knew what the context was before hearing the announcement*."

These responses reflect strategic and anticipatory listening behavior and align with best practices in listening instruction, where learners are encouraged to activate background knowledge and preview task content.

However, a large subset of participants reported minimal or passive strategies, including: *"Start listening"; "Concentrate and wait for the announcement"; "Wait for the audio?"* This suggests a lack of procedural clarity or training, particularly in using the preview window effectively. The response *"Wait for the audio?"* and comments such as *"Panic, haha, 30 seconds is not a lot of time!"* also reveal task anxiety or uncertainty, potentially linked to instruction wording or previous test experience. Notably, only a few participants mentioned note-taking or physically preparing themselves (e.g., grabbing a pen, finding a quiet space), suggesting that external test-taking routines are not widely established or may be assumed rather than consciously executed. In general, the data reveal a divide between strategic and reactive approaches to listening. This variation supports the argument that even well-formulated instructions cannot substitute for listening strategy training. Learners may comprehend the instructions but fail to activate efficient test-taking behaviors unless these behaviors have been explicitly taught and practiced. The findings underscore the need for integrated instruction that combines clear task input with strategic scaffolding – particularly in high-stakes contexts where every second counts.

When participants were asked which part of the instruction was most important to them, their responses clustered around three main themes: (1) Task procedure, (2) Timing and repetition, and (3) Contextual framing. Interestingly, despite being expressed in three

subtly different ways across instruction types, the detail that the recording would be played twice was consistently cited as important – even though few respondents commented on the difference in formulation. This suggests that some information is so critical that its mere presence matters more than its wording.

When participants were asked whether anything in the instruction was unclear or missing, most responded "No", suggesting a generally high level of surface comprehension. However, a minority flagged specific ambiguities, offering valuable insights into micro-level breakdowns in processing. One participant noted needing to look up the word "бланк", pointing to potential lexical barriers, especially for L2 users with limited academic or test-related vocabulary. This illustrates how single words can become friction points even in otherwise clear instructions, and underscores the importance of linguistic accessibility. There were several questions about task parameters and scoring logic, where participants highlighted missing or ambiguous information about response expectations: *"Is there only one correct answer or more?"; "What the questions actually are."*

Even if these elements are part of the task rather than the instruction, their mention suggests that test-takers often seek reassurance about the answer format before beginning. This may be especially true in high-stakes contexts, where clarity is not just about understanding, but about performance confidence.

Despite the information about the recording being played twice being presented in three different formulations across the instruction versions, participants rarely remarked on this detail. This suggests several possibilities: first, that this aspect is perceived as a fundamental or self-evident part of the listening task, thus participants did not feel the need to focus on it or elaborate on it in their feedback. Second, the variations in wording were semantically similar enough that changes in style or length did not significantly impact participants' perception or the importance they attributed to this information. Third, other elements of the instructions – such as the preview time or the communicative context – elicited stronger cognitive or emotional responses, relegating the "played twice" detail to a less salient position. Finally, it may indicate that this procedural information is automatically processed and does not require additional emphasis to support test-takers' understanding.

In summary, the findings suggest that optimal listening instructions should strike a balance between informativeness and cognitive economy. They should be sufficiently structured to support less experienced or anxious test-takers, while not overburdening advanced users. The results also indicate the potential value of differentiated or adaptive instructions, tailored to test-takers' proficiency levels, cognitive styles, and emotional profiles – particularly in high-stakes contexts where clarity and fairness are paramount. In conclusion, while the repetition of the recording is a critical procedural feature, its communicative effectiveness appears largely independent of stylistic variations in phrasing. For test-takers, this information is likely already an established component of the task format, reducing the need to highlight it explicitly in instructional texts.

Comparison of listening task instructions in five high-stakes tests (IELTS, TOEFL, TestDaF, Polish and Czech national exams) reveals both common approaches and significant differences, conditioned by the overall test concept and the type of communicative competence model underlying it.

All tests use high-frequency vocabulary, but IELTS and TOEFL sometimes include test terminology ("summary completion," "matching"), which may be challenging for less prepared candidates. Polish and Czech tests avoid such terms, using maximally

simple and unambiguous language. All tests predominantly use the imperative: short, clear instructions like "Choose the correct answer" or "Vyberte správnou možnost." IELTS and TestDaF provide audio instructions in a more natural style ("You will hear a conversation between…"), whereas TOEFL, Polish, and Czech instructions remain formally neutral. The communicative aspect is most often realized through organizational functions: instructions on the test start, number of tasks, listenings, etc. The metacognitive function (strategy orientation) is partially implemented through time indications for task familiarization. Contextualization is clearly present in IELTS and TestDaF but only partially or entirely absent in Polish and Czech tests. All tests show a high degree of formalization of written instructions. In particular, IELTS strictly regulates answer format ("Write NO MORE THAN THREE WORDS…"). Polish and Czech tests also provide clear technical instructions without excessive detail. TestDaF demonstrates the best balance: formalization combined with contextual clues. IELTS and TestDaF clearly embed tasks in realistic situations ("You are at a bus stop…"), enhancing comprehension and test authenticity. In the Czech test, context is fragmentary but functional ("Uslyšíte rozhovor s odborníkem…"), while the Polish test almost completely avoids situational framing.

This analysis confirms that the effectiveness of instructions depends not only on lexical or grammatical simplicity but also on the communicative balance between formalization, context, and cognitive accessibility.

Conclusions. This study demonstrates that task instructions in high-stakes listening assessments are not merely procedural add-ons but integral components of the test construct. They shape candidates' comprehension, strategy use, and emotional readiness, thereby influencing both performance outcomes and construct validity.

Across international exams such as IELTS, TOEFL, TestDaF, and national Slavic-language tests, several key tendencies emerge. The most effective instructions balance clarity, brevity, and contextual grounding. Overly minimalist formats risk ambiguity and insufficient cognitive orientation, whereas excessively detailed ones increase cognitive load and anxiety. The optimal model integrates a small number of essential elements – task type, response format, number of listenings, and brief contextual framing – while maintaining linguistic simplicity and procedural transparency.

The comparative analysis also reveals that communicative contextualization significantly enhances test-taker orientation, especially in lower-proficiency users. When instructions identify the setting, roles, or topic of the recording, candidates activate relevant background knowledge and engage in more strategic listening. Conversely, purely formal instructions, though standardized, can limit comprehension and reduce ecological validity.

From a linguistic perspective, imperative forms and high-frequency vocabulary are most effective in ensuring accessibility across proficiency levels. Grammatical simplicity and consistency help minimize construct-irrelevant variance caused by misunderstanding of instructions. Lexical and syntactic economy should therefore be treated as central design principles.

Empirical findings confirm that instruction preferences are shaped by proficiency, cognitive style, and anxiety sensitivity. Advanced and confident users tend to favor concise instructions, while lower-level or anxious test-takers prefer more structured and contextualized formats. This highlights the potential of adaptive instruction models, which could tailor guidance to candidate profiles or test complexity while maintaining fairness and standardization.

For the design of Ukrainian language proficiency tests, the findings indicate that effective listening task instructions should follow several key principles. They should maintain a clear organizational structure and a consistent format across all tasks to help test-takers quickly recognize familiar patterns and focus on comprehension rather than navigation. The instructions should include only the minimal but essential contextual information, ensuring clarity without overloading the test-taker with unnecessary details.

Linguistically, it is recommended to use simple, high-frequency verbs in the imperative mood, as they are easier to process and more universally understood by learners of various proficiency levels. At the same time, the use of complex syntax or culturally specific references should be avoided, as these may confuse or disadvantage non-native speakers.

Finally, the instructions should include a supportive and transparent framing of time constraints, helping to reduce test anxiety and promote better cognitive focus during the listening task. Together, these principles contribute to clearer, fairer, and more contextually valid assessment practices.

Ultimately, well-designed listening task instructions serve as both cognitive and emotional scaffolding. They help test-takers focus on communicative content rather than procedural uncertainty, contributing to more valid, reliable, and equitable assessment outcomes.

References

Theoretical and research sources

1. Alderson J. C., Clapham C., Wall D. Language Test Construction and Evaluation. Cambridge : Cambridge University Press, 1995. 310 p.
2. Bachman L. F., Palmer A. S. Language Testing in Practice: Designing and Developing Useful Language Tests. Oxford : Oxford University Press, 1996. 377 p.
3. Elliott M., Wilson J. Context validity. *Examining Listening: Research and Practice in Assessing Second Language Listening* / ed. by A. Geranpayeh, L. Taylor. Cambridge : UCLES, Cambridge University Press, 2013. P. 152–240.
4. Fulcher G. Practical Language Testing. London : Routledge, 2010. 370 p.
5. Green R. Designing Listening Tests: A Practical Approach. UK : Palgrave Macmillan, London, 2017. 210 p.
6. Meisuri, M. R. S. Language Testing: Foundations, Approaches, Innovations, and Future Direction. Publisher, umsu press, 2024. 260 p.
7. Shohamy E. The Power of Tests: A Critical Perspective on the Use of Language Tests. London : Longman, 2001. 190 p.
8. Taylor L., Geranpayeh A. Assessing listening for academic purposes: Defining and operationalising the test construct. *Journal of English for Academic Purposes.* 2011. Vol. 10. No. 2. P. 89–101.
9. Vandergrift L., Goh C. Teaching and Learning Second Language Listening: Metacognition in Action. New York : Routledge, 2012. 338 p.

Official standards and examination materials

1. Council of Europe. Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). Strasbourg : Council of Europe Publishing, 2001. 273 p.
2. Council of Europe. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume. Strasbourg : Council of Europe Publishing, 2020. 237 p.
3. Educational Testing Service (ETS). TOEFL iBT Test: Information and Registration Bulletin. Princeton, NJ : ETS, 2024. 56 p.

4. Cambridge Assessment English. IELTS Information for Candidates. Cambridge : Cambridge University Press, 2024. 32 p.

5. State Commission for the Certification of Proficiency in Polish as a Foreign Language. Państwowy Egzamin Certyfikatowy z Języka Polskiego jako Obcego: Informator. Warszawa : Ministerstwo Edukacji i Nauki, 2023. 74 p.

6. Univerzita Karlova. Ústav jazykové a odborné přípravy. Certifikovaná zkouška z češtiny pro cizince (CCE): Informace pro uchazeče. Praha : Univerzita Karlova, 2023. 45 s.

7. TestDaF-Institut. Test Deutsch als Fremdsprache (TestDaF): Handbuch. Bochum : TestDaF-Institut, 2024. 98 p.

ДО ІДЕАЛЬНОЇ ІНСТРУКЦІЇ З АУДІЮВАННЯ: ПОРІВНЯЛЬНИЙ АНАЛІЗ ВИСОКОСТАВКОВИХ ТЕСТІВ

Ксенія Бородін

Український католицький університет
Школа української мови і культури
вул. Свєнціцького, 17, 79000, Львів, Україна
тел. 097 822 66 75
ел. пошта: borodinka@ucu.edu.ua
https://orcid.org/0009-0001-9307-1363

У статті проаналізовано особливості формулювання інструкцій до завдань у субтесті з аудіювання у високоставкових іспитах, зокрема тестах з англійської мови IELTS та TOEFL, з німецької мови TestDaF, з чеської мови (Certifikovaná zkouška z češtiny pro cizince (CCE) та національному іспиті з польської мови (Państwowy Egzamin Certyfikatowy z Języka Polskiego jako Obcego). Порівняльний аналіз продемонстрував, що способи подачі інструкцій у міжнародних тестах істотно різняться за структурою, обсягом і рівнем комунікативної насиченості. Виявлено кілька основних моделей – від максимально лаконічних інструкцій, що зосереджуються лише на процедурній частині, до розгорнутих комунікативно орієнтованих форм, які спрямовані на створення автентичного контексту спілкування.

Окрему увагу приділено порівнянню змісту інструкцій із вимогами Загальноєвропейських рекомендацій з мовної освіти (CEFR). Зіставлення показало, що інструкції у більшості тестів радше зумовлені типом завдання, а не безпосередньо мовним рівнем, тому переважно не передають очікуваної когнітивної складності чи типу мовленнєвих дій, властивих відповідним рівням володіння мовою.

Для перевірки впливу мовного формулювання на розуміння тестових завдань було проведено власне дослідження на прикладі інструкцій українською мовою. У межах експерименту проаналізовано реакції та відповіді респондентів, що дало змогу виявити типові труднощі, пов'язані з когнітивним навантаженням, надмірною складністю або, навпаки, надмірним спрощенням інструкцій, яке може знижувати автентичність комунікативної ситуації.

Отримані результати підтверджують, що чіткість, структурна організація та мовна прозорість інструкцій безпосередньо впливають на якість виконання завдань з аудіювання. На основі аналізу сформульовано низку практичних рекомендацій для розробників тестів з української мови як іноземної. Зокрема, запропоновано принципи укладання інструкцій, які забезпечують баланс між доступністю, когнітивною оптимальністю та комунікативною релевантністю. Рекомендації мають на меті підвищення контекстуальної валідності, зменшення впливу нецільових факторів на результати оцінювання та підвищення надійності тестів загалом.

Ключові слова: аудіювання, інструкції до завдань, українська як іноземна, тестування, контекстуальна валідність.

Стаття надійшла до редакції 13.01.2026
доопрацьована 17.01.2026
прийнята до друку 22.01.2026