



ТЕОРІЯ ТА МЕТОДИКА НАВЧАННЯ

УДК 378.147:811.111'243:81'342:81'36

DOI <https://doi.org/10.5281/zenodo.20437751>

**English Beyond the Textbook: Grammar, Pronunciation and Sociocultural
Variation in Spoken Communication**

Violetta Panchenko,

PhD in Pedagogy, Associate Professor,

Associate Professor at the Department of Foreign Philology,

Municipal Establishment «Kharkiv Humanitarian Pedagogical Academy» of Kharkiv

Regional Council, 61001,7 Rustaveli Lane, Kharkiv, Ukraine

<https://orcid.org/0000-0002-2958-5802>

Iryna Shkola,

PhD in Philology, Associate Professor,

Head of the Department of Foreign Languages and Teaching Methods,

Berdyansk State Pedagogical University,

69000, 55A Universytetska Street, Zaporizhzhia, Ukraine,

<https://orcid.org/0000-0002-1455-6371>

Bohdana Saliuk,

PhD in Philology, Associate Professor,

Associate Professor at the Department of Foreign Languages and Teaching Methods,

Berdyansk State Pedagogical University,

69000, 55A Universytetska Street, Zaporizhzhia, Ukraine,

<https://orcid.org/0000-0002-9952-2399>

Accepted: 06.05.2026 | Published: 29.05.2026



Abstract: *The article addresses the problem of preparing learners to understand authentic spoken English, which often differs considerably from the controlled, standardised and textbook-based model of English commonly presented in EFL instruction. The **objective** of the study is to substantiate and model a pedagogical approach to teaching spoken English with attention to grammatical, pronunciation-related and sociocultural variation. The study employs descriptive-analytical, comparative, interpretive and pedagogical modelling **methods**. Its theoretical basis draws on recent research in EFL methodology, Global Englishes, sociolinguistics, phonetic awareness, spoken grammar and sociocultural competence. The research procedure includes conceptual analysis of scholarly literature, synthesis of methodological approaches to authentic listening instruction and the development of the Voice–Form–Context Framework. The **results** presented in the article suggest that authentic spoken English should not be treated as a deviation from a standard norm, but as a layered communicative event in which the speaker’s voice, the grammatical organisation of spontaneous speech and the sociocultural context interact. The study proposes a five-stage pedagogical sequence for working with authentic spoken material: First Encounter with a Voice, Language Clues in Context, Grammar in the Wild, Rebuilding the Message and Voice Profile. Each stage is aimed at developing learners’ receptive phonetic flexibility, awareness of spoken grammar, ability to interpret sociocultural meanings and capacity to apply listening strategies in conditions of linguistic variation. It is **concluded** that the proposed approach shifts EFL instruction from the reproduction of a single pronunciation or grammar norm towards the development of communicative flexibility, metalinguistic awareness and readiness to perceive English in its real spoken forms. The framework can be applied in secondary and tertiary EFL contexts, particularly in teacher education, where future English teachers need to learn how to select, interpret and pedagogically adapt authentic oral materials.*



Keywords: EFL instruction, authentic materials, oral discourse, language awareness, sociolinguistic competence, communicative flexibility, receptive skills.

**Англійська мова поза підручником: граматика, вимова й соціокультурна
варіативність усного мовлення**

Панченко Віолетта Володимирівна,

кандидат педагогічних наук, доцент,

доцент кафедри іноземної філології Комунального закладу «Харківська
гуманітарно-педагогічна академія» Харківської обласної ради,

61001, пров. Руставелі, 7, м. Харків, Україна

<https://orcid.org/0000-0002-2958-5802>

Школа Ірина Вікторівна,

кандидат філологічних наук, доцент,

завідувач кафедри іноземних мов і методики викладання Бердянського
державного педагогічного університету,

69000, вул. Університетська, 55А, м. Запоріжжя, Україна

<https://orcid.org/0000-0002-1455-6371>

Салюк Богдана Анатоліївна

кандидат філологічних наук, доцент, доцент кафедри іноземних мов і методики
викладання Бердянського державного педагогічного університету,

69000, вул. Університетська, 55А, м. Запоріжжя, Україна

<https://orcid.org/0000-0002-9952-2399>

*Анотація: Стаття присвячена проблемі підготовки здобувачів освіти до
розуміння автентичного усного англійського мовлення, яке істотно*



відрізняється від контрольованої, стандартизованої та переважно підруничкової моделі англійської мови. **Метою** дослідження є теоретичне обґрунтування й методичне моделювання підходу до навчання англійського усного мовлення з урахуванням граматичної, фонетичної та соціокультурної варіативності. У дослідженні було використано такі **методи**: описово-аналітичний, порівняльний, інтерпретаційний і методико-моделювальний підхід. Теоретичну основу становлять сучасні дослідження з методики навчання англійської мови як іноземної, *Global Englishes*, соціолінгвістики, фонетичної обізнаності, розмовної граматики та формування соціокультурної компетентності. Здійснено концептуальний аналіз наукових праць, синтез методичних підходів до навчання автентичного усного мовлення та розроблено авторську модель *Voice–Form–Context Framework*. Отримані результати дозволяють стверджувати, що реальне усне англійське мовлення доцільно розглядати не як відхилення від стандартної норми, а як багаторівневу комунікацію, у якій взаємодіють голос мовця, граматична організація спонтанного мовлення та соціокультурний контекст. Запропоновано п'ятиетапну методичну послідовність роботи з автентичним усним матеріалом: *First Encounter with a Voice, Language Clues in Context, Grammar in the Wild, Rebuilding the Message, Voice Profile*. Кожен етап спрямований на розвиток рецептивної фонетичної гнучкості, усвідомлення особливостей розмовної граматики, уміння інтерпретувати соціокультурні смисли та здатність застосовувати стратегії слухання в умовах мовної варіативності. Сформульовано висновок про те, що запропонований підхід зміщує акцент із відтворення єдиної вимовної або граматичної норми на формування комунікативної гнучкості, метамовної обізнаності та готовності до сприйняття англійської мови в її реальному усному вияві. Розроблену модель можна застосовувати в контексті викладання англійської мови як іноземної у середній та вищій школі, зокрема в педагогічній освіті, де майбутнім вчителям



англійської мови необхідно навчитися відбирати, інтерпретувати та педагогічно адаптувати автентичні усні матеріали.

Ключові слова: викладання англійської мови як іноземної, автентичні матеріали, розмовний дискурс, мовна обізнаність, соціолінгвістична компетентність, комунікативна гнучкість, рецептивні вміння.

Problem statement. In many EFL contexts, learners encounter English primarily as a carefully selected, pedagogically controlled and largely standardised language system. Coursebooks, classroom recordings and examination materials often present English through stable grammar patterns, clearly articulated pronunciation models and culturally neutral dialogues. Such materials are necessary at the early stages of language learning because they reduce cognitive load and provide learners with a manageable linguistic foundation. However, when this model remains dominant at intermediate and advanced levels, it may create a gap between classroom English and the English that learners hear in films, interviews, podcasts, academic mobility contexts, international projects, workplace communication and everyday digital interaction.

This gap is not limited to pronunciation. Authentic spoken English differs from textbook English at several interconnected levels. First, it is phonetically variable: speakers use different accents, speech rates, rhythmical patterns, reductions, connected speech processes and intonation contours. Second, it is grammatically dynamic: spontaneous speech contains ellipsis, hesitation, self-repair, discourse markers, vague language, fillers, reduced forms and unfinished syntactic structures. Third, it is socioculturally situated: meanings are shaped by region, social background, communicative situation, interpersonal distance, pragmatic conventions and culturally recognisable references. For this reason, the challenge facing EFL instruction is not simply to expose learners to ‘different accents’, but to teach them to process spoken English as a complex communicative event.



The problem becomes especially important in the context of contemporary language education, where learners increasingly access English through digital platforms rather than only through formal instruction. Streaming services, social media, online lectures, international webinars, video blogs and collaborative learning environments expose learners to English as a pluricentric and socially diverse language. At the same time, many learners still tend to interpret unfamiliar pronunciation, non-textbook grammar or regionally marked expressions as signs of incorrectness, low language quality or personal listening failure. This response demonstrates not only a linguistic difficulty, but also a pedagogical limitation: learners are insufficiently prepared to understand variation as a normal and meaningful feature of spoken communication.

Recent work in Global Englishes and English as a lingua franca pedagogy has questioned the continued dominance of a narrow native-speaker model in English language teaching. Rose, McKinley and Galloway emphasise that the global use of English requires pedagogical approaches that prepare learners for linguistic diversity rather than only for conformity to standard native-speaker norms [1]. Empirical studies of accent perception and listening also demonstrate that exposure to different English accents can support learners' multidialectal listening skills and help them become more flexible listeners [2]. These findings suggest that variation should be treated not as an additional cultural topic, but as a core dimension of communicative competence.

In Ukrainian secondary and tertiary education, this issue has additional relevance. Learners and students increasingly participate in international academic, professional and humanitarian communication, where English is used by speakers from different regions and linguistic backgrounds. Future English teachers, in particular, need not only to understand such variation themselves, but also to develop the methodological competence required to introduce authentic spoken materials responsibly and systematically. Therefore, EFL instruction should move beyond the exclusive reliance on textbook models and create pedagogical conditions for learners



to notice, interpret and gradually manage grammatical, pronunciation-related and sociocultural variation in spoken communication.

The present article addresses this need by proposing the Voice–Form–Context Framework, a pedagogical model for working with authentic spoken English in EFL instruction. The model is based on the assumption that learners' readiness for real communication depends on their ability to understand how voice, linguistic form and context interact in oral discourse. Voice refers to pronunciation, accent, rhythm, connected speech and intelligibility. Form refers to spoken grammar, including discourse markers, ellipsis, hedging, vague expressions, reduced forms and syntactic incompleteness typical of spontaneous speech. Context refers to sociocultural meanings, pragmatic intentions, regional identity and communicative situation. Taken together, these three dimensions provide a methodological basis for teaching English beyond the textbook while preserving a clear pedagogical structure.

Analysis of recent research and publications. The problem of preparing EFL learners for authentic spoken English has been increasingly addressed in recent research on Global Englishes, multidialectal listening, pronunciation instruction, spoken grammar, authentic materials and intercultural communicative competence. These areas are directly relevant to the present study because they question the sufficiency of textbook-based models of English and foreground the need to develop learners' ability to understand linguistic variation in real communication.

A substantial theoretical basis for rethinking English language teaching is provided by Rose, McKinley and Galloway [1], who review pedagogical research in Global Englishes and demonstrate that language teaching can no longer be limited to the reproduction of a narrow native-speaker model. The authors argue for the inclusion of diverse Englishes, critical reflection on native-speakerism and the development of learners' readiness for communication in multilingual and multicultural settings. For the present study, this position provides the general theoretical justification for moving from textbook English to real-world spoken English.



The pedagogical relevance of Global Englishes has also been demonstrated in classroom-based research. Da Costa and Rose [2] report on a quasi-experimental classroom intervention embedded in a general English course and show that systematic work with Global Englishes can influence school-aged learners' perceptions of English. It supports the assumption that learners' awareness of different Englishes should not be treated as an optional cultural supplement, but as a regular component of communicative language education.

A more specific contribution to the listening dimension is made by Jeong, Lindemann, Forsberg and Ribbeklint [3], whose study focuses not only on learners' understanding of unfamiliar accents, but also on attitudes towards speakers of different Englishes. The relationship between accent, perceived difficulty and actual understanding is further clarified by Verbeke and Simon [4]. Their study distinguishes comprehensibility, accentedness and intelligibility of native and non-native English speech and shows that accentedness and intelligibility should not be treated as identical phenomena. For EFL instruction, this means that the aim should not be to remove variation from listening input, but to train learners to process unfamiliar speech patterns more flexibly.

The need to reconsider pronunciation instruction in the context of World Englishes is also emphasised by Almusharraf [5]. The author examines university EFL instructors' perceptions and practices concerning pronunciation instruction and identifies a gap between the acknowledged importance of pronunciation and actual classroom practice, which proves that work with real-world pronunciation requires not only learner training, but also methodological support for teachers.

The issue of textbook English and authenticity is addressed from another perspective by Hui and Tenbrink [6], who compare how second language learners and first language speakers perceive conversations in English learning materials. Their study highlights the fact that pedagogical conversations may be interpreted differently



by learners and native speakers and that materials designed for learning purposes do not always correspond to expectations of natural spoken interaction.

Authentic listening materials have also been examined by Alamri [7], who reviews the benefits and challenges of using authentic materials for listening purposes in EFL contexts.

The grammatical dimension of spoken communication is supported by corpus-based and applied linguistic research into discourse markers and spoken interaction. Huang, Lin and Gráf [8] examine the development of discourse marker use across CEFR fluency levels in a learner corpus. These forms are often absent from traditional grammar teaching or are treated as informal additions, although learners regularly encounter them in real speech.

The sociocultural dimension of English language teaching is addressed in recent research on intercultural communicative competence. Iswandari and Ardi [9] provide a systematic review of studies on EFL teachers' and pre-service teachers' intercultural communicative competence. Their analysis identifies several recurring themes: teachers' cognition, ICC assessment, teacher development, ICC training and factors influencing intercultural competence.

A Ukrainian perspective on intercultural communicative competence is represented by Mykytenko, Fedorchuk, Ivasyuta, Hrynya and Kotlovskyi [10], who examine intercultural communicative competence development in journalism students. Although their study is situated in ESP rather than EFL teacher education, it is valuable because it treats intercultural competence as an integral part of professional language training and combines theoretical conceptualisation with empirical data from Ukrainian higher education.

The competence-based orientation of the present article is also supported by the CEFR Companion Volume [11], which foregrounds plurilingual and pluricultural competence, mediation and flexible language use across contexts.



Earlier theoretical studies remain relevant for clarifying the conceptual basis of spoken grammar and language variation. Carter and McCarthy [12] argue that spoken grammar has its own regularities and should not be viewed as an incomplete or incorrect version of written grammar. Coupland [13] explains how linguistic variation functions as a marker of style and identity, while Lippi-Green [14] demonstrates how accent is connected with ideology, social evaluation and discrimination.

Identification of previously unresolved aspects of the general problem. The reviewed studies confirm the need to prepare EFL learners for linguistic and sociocultural variation in spoken English. However, several aspects of this problem remain insufficiently addressed in pedagogical research and classroom practice.

First, Global Englishes research has convincingly challenged the dominance of native-speaker norms, yet its classroom implementation often remains broadly conceptual. Teachers may accept the importance of linguistic diversity but still lack a clear procedure for integrating it into regular work with spoken materials.

Second, pronunciation, spoken grammar and sociocultural meaning are often examined separately. Accent studies focus mainly on intelligibility and speech perception; spoken grammar research describes features of oral discourse; intercultural studies analyse sociocultural competence. In authentic communication, however, these dimensions operate simultaneously and should therefore be taught in an integrated way.

Third, authentic listening materials are widely recommended, but their pedagogical use may be ineffective without systematic scaffolding. Learners need support in noticing how unfamiliar pronunciation, reduced grammatical forms, discourse markers, regional expressions and contextual clues contribute to meaning.

Purpose and objectives of the article. The purpose of the article is to theoretically substantiate and methodologically model an integrated approach to teaching authentic spoken English in EFL instruction through the interaction of pronunciation, spoken grammar and sociocultural meaning.

To achieve this purpose, the article sets the following objectives:



1. to analyse recent research on Global Englishes, multidialectal listening, pronunciation instruction, spoken grammar, authentic listening materials and intercultural communicative competence in relation to EFL pedagogy;
2. to clarify why authentic spoken English should be treated as a layered communicative event rather than as a listening text containing only lexical or phonetic difficulties;
3. to develop the Voice–Form–Context Framework as a methodological model for integrating pronunciation-related, grammatical and sociocultural dimensions of spoken communication;
4. to design a five-stage pedagogical sequence for working with authentic spoken English in secondary and tertiary EFL contexts;
5. to define the pedagogical value of the proposed approach for developing learners' receptive phonetic flexibility, spoken grammar awareness, sociolinguistic sensitivity and strategic listening competence.

Presentation of the main research material. The methodological premise of the present study is that authentic spoken English should not be reduced to a listening text used for checking comprehension. In real communication, spoken English functions as a layered communicative event in which pronunciation, grammar and sociocultural meaning interact dynamically. Learners hear not only words and sentences, but also voices, accents, rhythm, hesitation, discourse markers, pragmatic signals, regional references, social positioning and culturally shaped ways of speaking. Therefore, the pedagogical task is not merely to make learners understand a particular audio fragment, but to help them develop strategies for interpreting spoken communication beyond textbook models.

This position is consistent with the broader shift from a native-speaker-oriented model of English language teaching to approaches informed by Global Englishes and English as a lingua franca. Rose, McKinley and Galloway argue that English language education should prepare learners for communication across diverse linguistic and



cultural settings rather than for imitation of a single native-speaker norm [1]. Jenkins also emphasises that in lingua franca communication the central issue is not conformity to a prestigious accent, but intelligibility, attitude and successful negotiation of meaning [16]. For EFL pedagogy, this means that learners should be guided to perceive English variation as normal, meaningful and communicatively relevant.

At the same time, the movement beyond textbook English should not be understood as a rejection of standard language instruction. Standard grammar, controlled pronunciation models and structured classroom input remain necessary at the earlier stages of learning. The problem arises when such models are treated as the only legitimate form of English and when learners are not prepared for the variability of real spoken interaction. Authentic oral communication includes regional and social accents, connected speech, reduced forms, discourse markers, ellipsis, vague language and culturally situated meanings. These features may cause comprehension difficulties, but they are not accidental deviations. They are systematic elements of spoken discourse.

The article therefore proposes the Voice–Form–Context Framework as a methodological model for teaching authentic spoken English in EFL contexts. The framework is based on three interrelated dimensions: *Voice*, *Form* and *Context*.

Voice refers to the sound dimension of spoken English. It includes pronunciation, accent, rhythm, intonation, speech rate, connected speech, reductions and the degree of perceived intelligibility. Research on accent perception shows that accentedness, comprehensibility and intelligibility should not be treated as identical constructs [4]. A speaker may sound unfamiliar or strongly accented to learners, but the message can still be intelligible if learners are trained to attend to recurring phonetic patterns and contextual support. Previous Ukrainian research has also shown that regional accents can be used as a pedagogical resource for developing phonetic skills and listening flexibility [16]. In the present model, however, accent work is not treated as an isolated



phonetic topic. It becomes part of a broader process of learning to interpret spoken communication.

The notion of receptive phonetic flexibility is central to this dimension. Learners do not need to imitate regional accents or reproduce all phonetic features they encounter. Instead, they need to recognise that familiar words may sound different depending on accent, speech rate, reduction and connected speech. High variability phonetic training research supports this principle by showing that exposure to multiple talkers and varied phonetic contexts can improve second language speech perception [17]. Although the present framework is not designed as laboratory-based phonetic training, it draws on the same pedagogical logic: learners become better listeners when they hear different voices and learn to identify patterns across variation.

Form refers to the grammatical organisation of spontaneous spoken discourse. In many EFL classrooms, grammar is still presented mainly through written, complete and carefully structured sentences. Authentic spoken English, however, often contains incomplete clauses, ellipsis, self-repair, repetitions, short responses, vague expressions, discourse markers, fillers and hedging. Carter and McCarthy argue that spoken grammar should not be viewed as an inferior or incomplete version of written grammar, since it has its own regularities and communicative functions [12]. For learners, this distinction is essential. If they expect spoken English to follow the same patterns as written textbook examples, they may misinterpret natural speech as chaotic, incorrect or grammatically deficient.

The Form dimension therefore aims to develop spoken grammar awareness. Learners should be able to notice how speakers organise meaning in real time, manage hesitation, maintain interaction, soften statements, signal agreement or disagreement, and build coherence through discourse markers. Huang, Lin and Gráf demonstrate that discourse marker use develops across CEFR fluency levels and is closely related to spoken interaction [8]. In classroom practice, this means that forms such as *well, you know, I mean, sort of, actually, right, okay* and like should not be treated merely as



unnecessary fillers. They help speakers organise talk, manage interpersonal relations and guide listeners through spontaneous speech.

Context refers to the sociocultural and pragmatic dimension of oral communication. Spoken English is always situated: it reflects who is speaking, where the interaction takes place, what relationship exists between speakers, what cultural references are shared, and what pragmatic norms regulate the exchange. Coupland's work on language variation and identity shows that linguistic choices index social meanings and are connected with style, identity and social positioning [13]. Lippi-Green further demonstrates that accent is frequently evaluated through language ideology and social prejudice [14]. Therefore, teaching learners to understand spoken English also means helping them avoid simplistic or stereotyped interpretations of voices, accents and informal speech.

This dimension is closely related to intercultural communicative competence. Iswandari and Ardi show that intercultural competence in EFL settings is strongly connected with teacher cognition, teacher development and classroom practice [9]. Mykytenko, Fedorchuk, Ivasyuta, Hrynya and Kotlovskyi also demonstrate the relevance of intercultural communicative competence in Ukrainian higher education, particularly in professional language training [10]. Huang's study of EFL learners in Taiwan confirms that intercultural communicative competence can be developed through explicit classroom instruction [18]. These findings support the view that sociocultural interpretation should be integrated into listening work rather than taught only through separate culture-based topics.

The three dimensions of the framework can be summarised as follows.

Table 1

The Voice–Form–Context Framework for Teaching Authentic Spoken English

Dimension	Main focus	Typical features	Pedagogical aim
-----------	------------	------------------	-----------------



Voice	Pronunciation and sound variation	Accent, rhythm, intonation, connected speech, reduction, speech rate	To develop receptive phonetic flexibility and tolerance of unfamiliar voices
Form	Spoken grammar and discourse organisation	Ellipsis, discourse markers, fillers, hedging, vague language, self-repair, reduced forms	To develop awareness of how spontaneous spoken grammar works
Context	Sociocultural and pragmatic meaning	Regional references, communicative situation, identity, politeness, humour, pragmatic routines	To develop sociolinguistic sensitivity and contextual interpretation

Source: developed by the authors

The framework is not intended as a separate course module. Its main advantage is that it can be integrated into ordinary EFL instruction through carefully selected authentic spoken materials and structured classroom procedures. The following five-stage pedagogical sequence demonstrates how the framework may be applied in secondary and tertiary EFL contexts.

Stage 1. First Encounter with a Voice

The first stage is designed to capture learners' initial response to authentic spoken input. Learners listen to a short extract without prior explanation. The extract may be taken from an interview, vlog, podcast, public talk, informal conversation, documentary fragment or other authentic source. It should be short enough to avoid cognitive overload, usually between 30 and 90 seconds, and should contain a manageable number of unfamiliar features.

The task is not to identify the speaker's accent or translate the whole fragment. Learners are asked to record three things: what they understood, what they found difficult, and how the speaker's voice sounded to them. The purpose is to make learners aware of their first listening reaction. They may notice that the speaker speaks quickly,



reduces words, uses unfamiliar expressions, sounds regional, pauses often, repeats words or uses informal grammar. Such observations become the starting point for guided analysis.

This stage is important because learners often interpret difficulty in listening as personal failure. They may think that they «do not know English well enough» when, in fact, the problem may be unfamiliar accent, reduced forms, discourse density or cultural references. By discussing first impressions openly, the teacher helps learners understand that listening difficulty can be analysed and managed. The stage also creates an opportunity to address accent attitudes. Following Jenkins' argument about attitude and identity in English as a lingua franca [15], learners should be encouraged to see unfamiliar speech not as bad English, but as a legitimate form of communication that requires strategic listening.

Stage 2. Language Clues in Context

The second stage shifts attention from general difficulty to contextual inference. Learners listen again and identify clues that support comprehension. These clues may include repeated words, key lexical items, intonation, gestures or visual information if the material is video-based, speaker relationship, setting, topic, emotional tone and culturally recognisable references. The teacher may provide a guiding worksheet with questions such as: *Who seems to be speaking? What is the situation? What is the speaker trying to do? Which words or expressions helped you understand the topic? Which parts remained unclear?*

This stage develops the *Context* dimension of the framework. Learners are trained to treat listening as interpretation rather than decoding. They learn that meaning is not located only in separate words. It is created through the interaction of linguistic forms, situation, speaker intention and shared knowledge. This is especially important when learners encounter culturally marked expressions, regional phrases or pragmatic routines that do not have direct equivalents in textbook dialogues.



The stage also supports intercultural communicative competence. Studies on ICC in EFL settings confirm that intercultural competence develops more effectively when learners engage with meaning, perspective and communicative context rather than memorising cultural facts [9; 18]. In the proposed sequence, sociocultural content is not added after listening; it is embedded in the process of understanding the spoken text.

Stage 3. Grammar in the Wild

The third stage focuses on spoken grammar. Learners receive a short transcript of the same extract. The transcript should be pedagogically adapted only to the extent necessary for readability; it should preserve important features of spoken discourse, such as hesitations, repetitions, discourse markers, contractions, incomplete utterances and self-repair. The teacher asks learners to identify forms that would be unusual in formal written English but are natural in speech.

Examples may include discourse markers such as *well, you know, actually, I mean*; fillers such as *erm* and *uh*; vague expressions such as *sort of, kind of, or something*; ellipsis in short answers; reduced forms such as *gonna, wanna, dunno*; and incomplete or repaired sentences. The aim is not to teach learners to reproduce all of these features actively. Rather, the aim is to help them recognise that spontaneous spoken English is grammatically organised in a different way from written or scripted English.

This stage is grounded in spoken grammar research. Carter and McCarthy emphasise that spoken grammar has its own systematic patterns and communicative functions [12]. Huang, Lin and Gráf's corpus-based study of discourse markers across CEFR levels also suggests that these elements are linked to fluency and interactional competence [8]. Therefore, including spoken grammar in listening instruction helps learners move from the expectation of perfectly formed textbook sentences to a more realistic understanding of oral communication.



For future English teachers, this stage has particular methodological value. It teaches them not to correct every feature of spoken discourse as an error and not to simplify authentic speech to the point where it loses its natural character. Instead, they learn to select features that are pedagogically useful and explain them as functional elements of communication.

Stage 4. Rebuilding the Message

The fourth stage is aimed at making comprehension development visible. After working with voice, contextual clues and spoken grammar, learners listen to the extract again and reconstruct the message more accurately. They may complete a summary, arrange key ideas in order, compare their first and second understanding, or answer comprehension questions that require attention to both content and pragmatic meaning.

The key methodological principle of this stage is comparison. Learners should see what changed between the first listening and the later listening after guided noticing. They may realise that a phrase they initially missed was a reduced form of a familiar expression, that a discourse marker signalled a shift in meaning, that a regional word was understandable from context, or that an unfamiliar accent became clearer after attention was drawn to recurring sound patterns.

This stage corresponds to the pedagogical logic of authentic materials research. Alamri notes that authentic materials can be beneficial for listening development, but their effectiveness depends on careful selection, scaffolding and task design [7]. In the proposed sequence, learners are not simply exposed to difficult input. They are guided through stages that transform authentic listening into a reflective and manageable process.

The stage also connects with findings from Global Englishes classroom-based interventions. Da Costa and Rose show that classroom innovation can influence learners' perceptions of English and help them understand variation more positively [2]. Similarly, Jeong, Lindemann, Forsberg and Ribbeklint demonstrate that guided contact with unfamiliar Global Englishes accents can address both understanding and



attitudes [3]. In the *Voice–Form–Context* sequence, repeated listening after guided analysis helps learners experience variation not as an obstacle, but as a feature that can be interpreted.

Stage 5. Voice Profile

The final stage requires learners to produce a short analytical profile of the spoken extract. This profile should not be a stereotyped identification of the speaker's nationality or social background. Instead, it should describe observable features of the extract within the three dimensions of the framework.

A basic Voice Profile may include the following prompts:

What features of the speaker's voice did you notice?

Which pronunciation features or connected speech patterns affected your understanding?

Which spoken grammar features appeared in the extract?

Which discourse markers, fillers, reduced forms or incomplete structures were used?

What contextual or sociocultural clues helped you interpret the message?

What listening strategies were useful?

What became clearer after repeated listening and guided analysis?

For pre-service English teachers, the task can be extended into pedagogical design. Students may be asked to select a short authentic extract, prepare a *Voice–Form–Context* analysis, design two or three scaffolding tasks and explain how they would use the material with learners of a particular level. This turns the task from receptive analysis into methodological training.

The Voice Profile stage is important because it consolidates learners' metalinguistic awareness. They learn to describe spoken English professionally and respectfully, avoiding judgments such as 'wrong', 'bad', 'strange' or 'incorrect'. Instead, they use analytical categories: accent feature, discourse marker, reduced form, contextual clue, pragmatic meaning, listening strategy. This is particularly relevant in



teacher education, where future teachers need a language for discussing authentic speech without reinforcing accent prejudice or simplistic ideas of correctness.

The five-stage sequence can be used flexibly. At lower levels, teachers may work with very short extracts, provide visual support, pre-teach key vocabulary and focus on one or two features. At higher levels, learners may analyse longer extracts, compare different speakers, examine transcripts in greater detail and design their own tasks. In both cases, the pedagogical goal remains the same: to help learners move beyond the expectation of fully standardised textbook English and develop readiness for real spoken communication.

The proposed framework also has implications for assessment. Traditional listening tests often focus on correct answers after exposure to carefully selected or standardised audio. Such tasks are useful for measuring certain aspects of comprehension, but they do not fully assess learners' ability to manage variation. In the *Voice–Form–Context* approach, assessment may include reflective listening logs, comparison of first and second comprehension, short analytical commentaries, learner-generated glossaries of spoken features, mini-portfolios of authentic listening extracts and micro-teaching tasks for pre-service teachers.

Such assessment formats are aligned with the CEFR Companion Volume, which foregrounds plurilingual and pluricultural competence, mediation and flexible language use across contexts [11]. They also reflect the need to evaluate not only whether learners understood a particular text, but how they approached difficulty, which strategies they used, and whether they became more aware of the relationship between voice, form and context.

The pedagogical value of the framework lies in its integrative nature. It does not isolate pronunciation from grammar or culture. It treats spoken English as a communicative whole. *Voice* helps learners understand how speech sounds; *Form* helps them understand how spontaneous grammar works; *Context* helps them understand why the message is shaped in a particular way. Together, these dimensions



create a realistic model of oral communication and provide teachers with a practical structure for working with authentic spoken materials.

For Ukrainian EFL education, this model may be particularly useful in two contexts. In secondary education, it can help learners develop listening flexibility and reduce anxiety when encountering unfamiliar voices in digital media, international exams, mobility contexts or online communication. In higher education, especially in English teacher education, it can support future teachers in selecting authentic materials, analysing spoken discourse and designing tasks that prepare learners for real-world communication. Thus, the model contributes not only to listening instruction, but also to the development of methodological competence.

The transition from textbook English to authentic spoken communication should be gradual, guided and pedagogically purposeful. Learners need standard models, but they also need to understand that English exists beyond those models. When grammar, pronunciation and sociocultural variation are taught together, learners become better equipped to interpret real speech, tolerate unfamiliarity and participate more confidently in communication across diverse English-speaking contexts.

Conclusions. The purpose of the article was to theoretically substantiate and methodologically model an integrated approach to teaching authentic spoken English in EFL instruction through the interaction of pronunciation, spoken grammar and sociocultural meaning. The analysis has shown that the gap between textbook English and real spoken communication is not limited to unfamiliar accents or fast speech. It concerns a broader set of features, including pronunciation variability, reduced and interactional grammatical forms, discourse markers, pragmatic routines, regional expressions and culturally situated meanings.

The proposed *Voice–Form–Context Framework* addresses this gap by treating authentic spoken English as a layered communicative event. The *Voice* dimension draws learners' attention to pronunciation, accent, rhythm, intonation, connected speech and intelligibility. The *Form* dimension helps them recognise spoken grammar



as a functional system that includes ellipsis, fillers, hedging, vague language, discourse markers, reduced forms and self-repair. The *Context* dimension develops learners' ability to interpret sociocultural and pragmatic meanings embedded in oral communication.

The five-stage pedagogical sequence – *First Encounter with a Voice, Language Clues in Context, Grammar in the Wild, Rebuilding the Message and Voice Profile* – provides a practical structure for integrating authentic spoken materials into secondary and tertiary EFL instruction. The sequence enables learners to move from initial listening difficulty to guided noticing, contextual interpretation, metalinguistic reflection and more confident reconstruction of meaning. It also offers future English teachers a methodological tool for selecting, analysing and adapting authentic oral materials without reducing them to isolated pronunciation or vocabulary exercises.

The study confirms that preparing learners for real spoken English requires a shift from the reproduction of a single standard model to the development of communicative flexibility, receptive phonetic awareness, spoken grammar awareness and sociolinguistic sensitivity. Such an approach does not reject standard language instruction, but complements it by helping learners understand how English functions in authentic interaction.

Further research may focus on empirical testing of the *Voice–Form–Context Framework* in different educational settings, including upper secondary school, university EFL courses and English teacher education programmes. Particular attention should be given to learners' attitudes towards unfamiliar spoken varieties, changes in listening comprehension after guided exposure to authentic materials, and the development of assessment tools for measuring progress in interpreting voice, form and context in spoken communication.

References



1. Rose H., McKinley J., Galloway N. Global Englishes and language teaching: A review of pedagogical research. *Language Teaching*. 2021. Vol. 54, Is. 2. P. 157–189. DOI: <https://doi.org/10.1017/S0261444820000518>
2. da Costa N., Rose H. The impact of Global Englishes classroom-based innovation on school-aged language learners' perceptions of English: An exercise in practitioner and researcher partnership. *System*. 2024. Vol. 121. Article 103263. DOI: <https://doi.org/10.1016/j.system.2024.103263>
3. Jeong H., Lindemann S., Forsberg J., Ribbeklint H. Fostering Global Englishes listeners: Addressing understanding and attitudes at a Swedish upper-secondary school. *TESOL Quarterly*. 2025. Vol. 59, № 4: P. 2234–2264. DOI: <https://doi.org/10.1002/tesq.70007>
4. Verbeke G., Simon E. Listening to accents: Comprehensibility, accentedness and intelligibility of native and non-native English speech. *Lingua*. 2023. Vol. 292. Article 103572. DOI: <https://doi.org/10.1016/j.lingua.2023.103572>
5. Almusharraf A. Pronunciation instruction in the context of world English: Exploring university EFL instructors' perceptions and practices. *Humanities and Social Sciences Communications*. 2024. Vol. 11. Article 847. DOI: <https://doi.org/10.1057/s41599-024-03365-y>
6. Hui Y., Tenbrink T. Second language learners' perceptions of conversations in English materials. *SAGE Open*. 2025. Vol. 15, Is. 2. DOI: <https://doi.org/10.1177/21582440251339659>
7. Alamri W. Evaluating the benefits and challenges of using authentic materials in EFL context for listening purpose. *Frontiers in Education*. 2025. Vol. 10. Article 1611308. DOI: <https://doi.org/10.3389/feduc.2025.1611308>
8. Huang L.-F., Lin Y.-L., Gráf T. Development of the use of discourse markers across different fluency levels of CEFR: A learner corpus analysis.



Pragmatics. 2023. Vol. 33, Is. 1. P. 49–77. DOI:
<https://doi.org/10.1075/prag.21016.hua>

9. Iswandari Y. A., Ardi P. Intercultural communicative competence in EFL setting: A systematic review. *rEFLECTIONS*. 2022. Vol. 29, Is. 2. P. 361–380. DOI:
<https://doi.org/10.61508/refl.v29i2.260249>

10. Mykytenko N., Fedorchuk M., Ivasyuta O., Hrynya N., Kotlovskyi A. Intercultural communicative competence development in journalism students. *Advanced Education*. 2022. Vol. 9, Is. 20. P. 121–131. DOI:
<https://doi.org/10.20535/2410-8286.261521>

11. Council of Europe. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume. Strasbourg : Council of Europe Publishing, 2020. 278 p.

12. Carter R., McCarthy M. Spoken grammar: Where are we and where are we going? *Applied Linguistics*. 2017. Vol. 38, Is. 1. P. 1–20. DOI:
<https://doi.org/10.1093/applin/amu080>

13. Coupland N. Style: Language Variation and Identity. Cambridge : Cambridge University Press, 2007. 220 p.

14. Lippi-Green R. English with an Accent: Language, Ideology and Discrimination in the United States. 2nd ed. London ; New York : Routledge, 2012. 384 p.

15. Jenkins J. English as a Lingua Franca: Attitude and Identity. Oxford : Oxford University Press, 2007. 284 p.

16. Panchenko V., Shkola I., Dubrova O. Innovative approaches to listening: Regional accents as a means of developing phonetic skills. *Інноваційна педагогіка*. 2025. № 83(1). С. 51–55. DOI: <https://doi.org/10.32782/2663-6085/2025/83.1.9>

17. Uchihara T., Karas M., Thomson R. I. High variability phonetic training (HVPT): A meta-analysis of L2 perceptual training studies. *Studies in Second*



Language Acquisition. 2025. Vol. 47, Is. 3. P. 794–827. DOI:
<https://doi.org/10.1017/S0272263125100879>

18. Huang L.-J. D. Developing intercultural communicative competence in foreign language classrooms: A study of EFL learners in Taiwan. *International Journal of Intercultural Relations*. 2021. Vol. 83. P. 55–66. DOI:
<https://doi.org/10.1016/j.ijintrel.2021.04.015>