

MULTIMODALITY AND AUTHENTIC ORAL COMMUNICATION IN LANGUAGE TEACHING

Israilov Jakhongirkhon Djamshidkhonovich,

Tashkent State Medical University

j.d.israilov@gmail.com

ORCID: <https://orcid.org/0009-0008-8193-795X>

Abstract: Authentic spoken interaction is widely valued in English language teaching, yet many teachers find it challenging to work with the multimodal complexity of real communication in classroom settings. While authenticity has been extensively discussed in general ELT, less attention has been paid to how learners can engage productively with interaction in which meaning is shaped not only by linguistic choices but also by tone, gesture and interactional organisation. Drawing on classroom experience, this article explores practical ways of helping learners notice linguistic and non-linguistic cues in short extracts of authentic spoken interaction. Using a medical context as an illustrative case, it shows how carefully structured classroom work can make multimodal input manageable without oversimplifying it. Teachers reported increased learner confidence and awareness of how speakers shape meaning in real time, with emerging effects on students' own classroom interaction. The article offers realistic ways of integrating authentic spoken interaction into classes.

Keywords: authentic spoken interaction; multimodal communication; linguistic and non-linguistic cues; classroom practice; spoken language awareness; communicative competence

Introduction

Authentic communication has long been viewed as a desirable yet challenging resource in language teaching. In English language teaching more broadly, questions of authenticity have been explored in relation to materials, classroom practices and learner engagement (e.g. Gilmore, 2007; Mishan, 2005). This work has highlighted the potential value of exposure to naturally occurring language, while also acknowledging the practical difficulties teachers face when working with authentic input in instructional settings. Much of this discussion, however, has focused on general or academic English, with less attention to the interactional complexity of spoken communication in professional contexts.

One setting in which this complexity becomes particularly visible is medical communication. Clinical interaction is inherently multimodal: meaning is conveyed not only through terminology and grammatical choices but also through tone, pacing, gesture, gaze and the organisation of activity within the consultation

(Heritage & Maynard, 2006; Candlin & Sarangi, 2011). For learners, this creates a layered interpretive challenge. Even when key vocabulary is recognisable, understanding how clinicians manage uncertainty, respond to patients' concerns or soften recommendations through prosody and stance requires sensitivity to cues that extend beyond lexis and grammar.

Interest in multimodality within applied linguistics has grown steadily (e.g. Kress & van Leeuwen, 2001; Norris, 2004), offering analytical tools for examining how meaning is constructed through multiple semiotic resources. However, the pedagogical implications of this work remain unevenly developed, particularly in relation to classroom practice. Teachers often express uncertainty about how to help learners engage with authentic spoken interaction without overwhelming them, and classroom materials frequently prioritise linguistic form while giving limited attention to how non-verbal and interactional features shape meaning. As a result, a gap can emerge between the types of spoken interaction learners encounter in classrooms and those they are expected to interpret beyond them.

The present article addresses this gap by examining a set of teaching practices designed to support learners' engagement with short extracts of authentic spoken interaction in a manageable and pedagogically principled way. Using a medical context as an illustrative case, the focus is on guiding learners to notice both linguistic and non-linguistic features of interaction and to explore how these features contribute to meaning in real time. Drawing on classroom experience and teachers' reflections, the article illustrates how structured attention to multimodal aspects of spoken interaction may support learner confidence and offers one possible way of extending approaches to authenticity within everyday ELT practice.

Overview of the approach and goals

The work described in this article grew out of a practical challenge repeatedly voiced by teachers: while learners may successfully memorise specialised terminology, they often struggle to make sense of how language is used in real spoken interaction. In authentic professional talk, meaning is shaped through hesitation, prosodic softening, overlapping turns, indirectness, and the careful balancing of authority and interpersonal sensitivity. These interactional features, well documented in studies of medical communication, are rarely visible in simplified classroom materials, leaving learners with limited opportunities to observe how reasoning and interpersonal work are realised through language in real time.

The teaching practices adopted here were developed to help learners engage with these aspects of spoken interaction in a manageable way. Rather than presenting full consultations, short extracts were selected that offered clear

examples of interactional features—such as modal expressions used to manage uncertainty, brief question sequences, reformulations designed to support understanding, and subtle shifts in tone or gesture. The intention was not to reproduce the full complexity of professional discourse, but to create space for guided attention to features that often remain unnoticed in classroom work with spoken language.

The work was organised around recurring cycles in which learners first encountered a short segment of authentic interaction, then examined selected linguistic and non-linguistic cues with teacher support, and finally experimented with these insights in brief classroom tasks. Teachers involved in this process noted that such cycles helped learners focus on specific communicative behaviours without feeling overwhelmed by the density of authentic spoken discourse.

Across its development, this way of working was refined in response to teacher observations and classroom experience. Its goal is to offer ELT teachers a practical means of introducing authentic spoken interaction into their classes in a form that supports noticing, reflection and communicative experimentation, while remaining compatible with the time constraints of regular courses. A medical context is used throughout as an illustrative case, but the underlying principles are intended to be transferable to other ELT settings where learners need support in interpreting real-world spoken communication.

Phased Implementation

The implementation follows seven phases. Linguistic and non-linguistic input is introduced progressively so that learners are not overloaded and can attend to specific features of spoken interaction in a focused way.

Phase 1: Material selection and preparation

Teachers select short extracts of authentic spoken interaction. Short segments tend to be more effective than extended recordings, as they allow learners to concentrate on particular interactional features without losing the overall communicative thread. In the medical context used here, suitable extracts often include brief diagnostic question sequences, instances of hedging (such as *It might be* or *Let's check first*), the use of modal verbs to express recommendation or uncertainty (*You should*, *You may need to*), specialised terminology embedded in interaction, and moments of reformulation when a speaker clarifies meaning for the listener.

Extracts that make non-verbal cues visible—such as nodding, hand gestures, changes in posture, or softening through tone—are particularly valuable, as they foreground the multimodal nature of spoken communication. At this stage, teachers prepare a small number of focused noticing prompts that draw learners'

attention to both linguistic choices and interactional behaviour, without attempting to cover all aspects of the extract.

Phase 2: Contextual orientation

Learners are briefly introduced to the communicative situation: who is speaking, what the immediate goal of the interaction is, and which stage of the exchange they are observing. This orientation helps learners understand why particular linguistic choices are made—for example, the use of indirect questions in sensitive moments or polite requests during physical examinations. Where relevant, teachers may draw attention to recurring interactional patterns, while avoiding an explicit theoretical explanation. The aim at this stage is to provide sufficient context to support interpretation, rather than to analyse the interaction in detail.

Phase 3: Initial exposure and guided noticing

Learners listen to or watch the extract once for overall meaning. The teacher then replays the episode and guides learners' attention toward selected linguistic features, such as how uncertainty is expressed, the use of softening devices (a little, a bit, possibly), compact question–answer sequences, or the replacement of specialised terminology with more accessible synonyms.

Alongside these linguistic elements, learners are encouraged to observe non-verbal and interactional cues, including gesture, intonation, body orientation and pauses. Attending to these features in parallel helps learners see how meaning is jointly shaped by language and behaviour in spoken interaction.

Phase 4: Focused reflection

Learners reflect on what they have noticed, either individually or through brief discussion. They may comment on why a speaker chooses a conditional form rather than a direct statement, or how changes in tone accompany the delivery of sensitive information. Attention is also given to how non-verbal cues support, reinforce or modify linguistic meaning. Short reflection tasks—such as completing a simple chart linking linguistic and non-linguistic features—help learners articulate these observations without requiring extensive metalanguage.

Phase 5: Communicative practice and adaptation

Learners engage in short communicative tasks that draw on the features they have observed. These tasks may include transforming direct statements into softened forms, practising stance markers (It seems like...), role-playing brief interactions using similar tone or gesture, producing short dialogues that embed terminology in context, or adjusting grammatical choices to express probability, reassurance or caution.

Rather than reproducing full consultations, these activities allow learners to experiment with selected linguistic and interactional resources in a controlled and focused way.

Phase 6: Feedback

Teachers provide targeted feedback on learners' use of both linguistic forms and non-verbal cues. This may involve commenting on the effective use of modal verbs, drawing attention to overly strong formulations, or noting gestures or intonation patterns that enhance clarity. Learners are encouraged to repeat the task or try alternative formulations, gradually developing greater flexibility and awareness in their spoken communication.

Phase 7: Integration into ongoing learning

Previously used extracts may be revisited, or new ones introduced, in order to highlight related interactional features. Learners compare how speakers adjust grammar, tone or terminology in response to different communicative demands. Over time, this supports a more sustained awareness of how authentic spoken interaction works and encourages learners to view linguistic and non-linguistic resources as interconnected aspects of communication rather than as separate skills.

Integrating into practice

The impact of working with authentic spoken interaction was explored through written reflections and informal feedback collected from teachers who implemented the approach as part of their regular classroom practice. Their comments pointed to three recurring observations, although not all were equally prominent across teaching contexts: learners' increased awareness of how spoken interaction functions in real settings, both linguistically and non-verbally; greater confidence and communicative flexibility in classroom interaction; and the practical value of the approach for teachers seeking workable ways of integrating authentic input into everyday lessons. Taken together, these themes suggest that structured engagement with authentic interaction can enrich spoken language work in ELT classrooms and support more realistic communicative preparation.

Developing communicative flexibility and learner confidence

One theme repeatedly highlighted by teachers was the increased confidence and adaptability shown by learners when working with spoken interaction. Teachers reported that learners began to use modal verbs, softening devices and stance markers more naturally after noticing them in authentic extracts. Rather than memorising fixed phrases, learners experimented with different ways of expressing uncertainty, reassurance or caution, adjusting their language in response to communicative context.

Several teachers described a visible change in learners' attitudes over time. Initially, many were hesitant to interpret tone or gesture, expressing concern about "getting it wrong." After repeated cycles of noticing and short communicative tasks, learners became more comfortable making decisions about linguistic and non-linguistic choices. As one teacher noted: "When they tried the mini-dialogue again, they adjusted the tone and phrasing themselves. They sounded much more like people engaged in real interaction, not just repeating lines."

This development is consistent with accounts of awareness-based and practice-oriented approaches that emphasise learners' ability to make informed communicative choices rather than to reproduce predetermined forms. As learners gained confidence, teachers also observed that they began transferring insights from one example to another, showing early signs of greater interactional sensitivity and emerging discursive competence.

What learners began to notice in real interaction

One of the most frequently mentioned outcomes was learners' growing understanding of how spoken interaction functions in real settings. Teachers noted that learners began to recognise not only specialised vocabulary, but also recurrent interactional patterns used to question, clarify, reassure or soften information. Several teachers described increased learner sensitivity to hedging, indirectness and short question-answer sequences—features that rarely appear in simplified classroom materials.

Teachers also reported that learners became more attentive to non-verbal aspects of interaction. They commented on tone, pauses, gesture and shifts in facial expression, particularly in moments where sensitive information was conveyed. As one teacher observed, learners "started paying attention to how the speaker looks at the listener or changes tone. They said it helped them understand the meaning even when the words were hard." This shift suggests a movement away from an exclusive focus on terminology toward a more situated understanding of how meaning is constructed through coordinated linguistic and non-linguistic resources.

Professional value for teachers: moving into classroom practice

Teachers consistently reported that the clearly structured phases made it easier to incorporate authentic spoken interaction into regular lessons without overwhelming either learners or lesson planning. The approach was described as "small and focused," allowing teachers to maintain clarity and classroom control while still engaging learners with real interaction.

Several teachers highlighted the value of explicitly linking linguistic features with non-verbal cues, an area they had previously found difficult to address pedagogically. As one teacher noted, "I always wanted to talk about tone and

gesture, but I wasn't sure how to build a task around it. The phases made it easier.” This suggests that the approach provided teachers with a workable framework for addressing aspects of spoken interaction that are often acknowledged as important but rarely operationalised in classroom practice.

Some instructors also reported that working with authentic interaction prompted them to reconsider how they select materials and design classroom tasks more generally. The approach was described as flexible, adaptable and compatible with existing curricula. Although illustrated here through a medical context, these observations point to the potential transferability of the underlying principles to other ELT settings where learners need support in interpreting authentic spoken communication.

Challenges

Despite the positive outcomes reported by teachers, several challenges were also identified. One recurring concern related to the availability of suitable authentic recordings. Teachers sometimes expressed uncertainty about whether they were selecting “the right moments” from longer interactions or whether a particular extract contained enough observable features to support focused noticing. Others noted that learners at lower proficiency levels required additional scaffolding in order to engage confidently with authentic spoken interaction.

Time constraints represented a further challenge. Teachers reported that a full noticing cycle occasionally extended beyond a single lesson, particularly when learners needed extra support with vocabulary or contextual understanding before engaging with the extract itself. A small number of teachers suggested that optional extension activities or shortened versions of the phases could be useful for classes operating under tighter time limitations.

Taken together, these challenges point to areas where additional pedagogical support could enhance the accessibility of the approach. Resources such as curated banks of short extracts, brief teacher guides, or ready-made noticing prompts may help reduce preparation time and support teachers who are less experienced in working with authentic spoken interaction.

Synthesising insights: toward a structured classroom strategy

Viewed collectively, the teacher reflections suggest that authentic spoken interaction can function as a practical and pedagogically valuable resource in ELT classrooms when introduced through a structured and learner-sensitive approach. Learners developed greater awareness of how spoken interaction operates, attending not only to linguistic forms but also to non-verbal and interactional cues. They also demonstrated increased confidence in making communicative choices, moving beyond the reproduction of isolated phrases toward more flexible use of

language in classroom interaction. Teachers, in turn, described the approach as manageable and adaptable across different teaching contexts.

These observations point to the potential scalability of the approach. Through the use of clearly sequenced phases, focused noticing activities and brief opportunities for communicative experimentation, teachers were able to integrate authentic spoken interaction into regular lessons without overwhelming learners. Although illustrated here through a medical context, the underlying strategy offers a replicable way for ELT practitioners to work with authentic spoken language in contexts where interactional complexity is high and classroom time is limited.

Discussion and wider implications

The findings of this project highlight the potential of providing learners with structured access to authentic spoken interaction, even when classroom time is limited and extracts are necessarily brief. What emerged most clearly from teachers' reflections is that learners were not simply exposed to engaging materials; rather, they began to develop a more nuanced understanding of how spoken interaction functions in real settings. Instead of approaching language use as a collection of discrete lexical items or as scripted exchanges, learners started to recognise that meaning is shaped through the coordinated use of language, tone, gesture and situational context.

This shift is significant because it resonates with long-standing observations in discourse and interactional research, which emphasise the multimodal nature of spoken communication and the role of indirectness, hesitation and prosodic softening in everyday meaning-making. However, such insights often remain distant from classroom practice, partly because teachers lack pedagogically workable ways of presenting authentic interaction without overwhelming learners. The phased sequence described here appears to offer one possible response to this challenge. By breaking complex interaction into observable components, it enabled teachers to guide learners' attention toward features that might otherwise go unnoticed in real-time communication.

One particularly interesting outcome was the way learners began to treat the extracts as material for interpretation rather than as listening tests to be completed successfully. This reframing appeared, at least in some classes, to shift learners' attention away from understanding every word and toward observing the choices speakers make as they manage uncertainty, negotiate interpersonal relations or adapt their language to situational demands. Teachers noted that learners commented on how speakers soften advice, rephrase information for clarity, or adjust intonation when moving between informational and interpersonal work. These observations suggest that authenticity, when pedagogically mediated, can

function not as an obstacle but as a resource for developing interactional awareness.

Such developments align with wider ELT discussions concerning the limitations of approaches that focus exclusively on either linguistic form or communicative fluency. Authenticity alone does not guarantee learning, just as theoretical principles do not translate into practice without careful scaffolding. The experience reported here suggests that a small-scale, iterative sequence—combining exposure, guided noticing, reflection and brief communicative experimentation—can make insights about multimodality and interactional pragmatics accessible within everyday ELT practice. In this respect, the approach contributes to ongoing debates about how real-world spoken communication can be integrated into the classroom in ways that remain pedagogically manageable.

Another important aspect of the findings is their potential transferability. Although the work was illustrated through a medical context, the challenges identified by teachers—bridging the gap between controlled practice and authentic interaction, supporting learners in interpreting tone and stance, and integrating authentic materials within limited classroom time—are familiar across many ELT settings. Teachers working in other professional domains, as well as those focusing on spoken fluency in general ELT courses, often face similar concerns. The fact that teachers with varying levels of experience described the sequence as manageable suggests that the underlying principles may be relevant beyond this specific context.

The challenges identified by teachers also point toward directions for further development rather than fundamental limitations of the approach. Difficulties in sourcing suitable authentic recordings, for example, are widely acknowledged in ELT, particularly in contexts where ethical considerations apply. This highlights the need for institutional collaboration and shared resource development. Similarly, the additional scaffolding required by lower-proficiency learners suggests that multimodal awareness may develop gradually and could be supported through differentiated tasks or carefully sequenced materials. These observations open avenues for future work on how multimodal spoken interaction can be integrated across proficiency levels.

Perhaps most striking is the way teachers described the process as a form of professional development for themselves as well as for their learners. Several noted that working with authentic spoken interaction prompted them to reconsider what counts as “teachable language” and to experiment with bringing tone, gesture and stance into classroom work alongside grammar and vocabulary. This underscores an important dimension of ELT innovation: pedagogical approaches need to be sustainable and meaningful for teachers, not only effective for learners. By encouraging teachers to engage analytically with spoken interaction alongside

their students, the approach described here has the potential to support more reflective and responsive classroom practices.

Overall, the findings suggest that authenticity can be made pedagogically accessible without simplifying away the complexity that characterises real communication. Through small but purposeful acts of guided interpretation, learners can begin to engage with spoken interaction as a multilayered social practice rather than as a purely linguistic task. While the medical context provides a particularly clear illustration of these dynamics, the broader implication is that ELT practice benefits from approaches that foreground how meaning is constructed through coordinated linguistic and non-linguistic resources. It is in this direction that current ELT research and pedagogy continue to move, and the work presented here offers one classroom-based illustration of how such principles can be realised in practice.

Conclusion

The experience described in this article demonstrates that authentic spoken interaction can be brought into the ELT classroom in ways that are both pedagogically manageable and instructionally meaningful. Rather than positioning authenticity as an overwhelming or advanced requirement, the work shows that short, carefully selected extracts of real interaction can serve as productive entry points for helping learners understand how spoken communication unfolds in real settings. When learners attend simultaneously to language, gesture, tone, pacing and pragmatic choices, they begin to view spoken interaction not as a fixed set of expressions to memorise, but as a dynamic process in which meaning is negotiated moment by moment.

A central insight emerging from this work is that multimodal awareness does not develop automatically. Many learners initially approached authentic recordings with the expectation that their task was to extract key words or identify correct answers. Only after repeated cycles of guided noticing did many learners begin to recognise how speakers soften advice, frame uncertainty or adjust their bodily behaviour to manage sensitive moments. These observations resonate with broader discussions in ELT and applied linguistics suggesting that exposure alone is rarely sufficient; learners benefit from structured opportunities to interpret, question and test their assumptions about communicative behaviour. In this respect, the phased sequence helped bridge the gap between abstract knowledge about communication and the realities of spoken interaction.

Teacher reflections also underscore the importance of classroom ecology in shaping how authenticity is approached. The process encouraged teachers to adopt a more exploratory stance, allowing space for learners to draw attention to features

that might otherwise go unnoticed. Several teachers noted that revisiting familiar materials through a multimodal lens prompted them to reconsider what counts as “teachable language” and to experiment with bringing tone, gesture and stance into classroom work alongside grammar and vocabulary. This suggests that the approach supported not only learner development but also teacher learning, highlighting the potential of classroom-based inquiry as a form of professional development.

Importantly, the practical nature of the work gives it relevance beyond the immediate illustrative context. Although the extracts were drawn from medical settings, the underlying principles—working with short stretches of authentic interaction, integrating linguistic and non-linguistic noticing, and using brief communicative tasks to test emerging insights—are applicable across a wide range of ELT contexts. Teachers working with spoken interaction in general ELT courses, as well as in other professional or academic domains, face similar challenges in helping learners interpret meaning under time pressure or in socially sensitive situations. The framework outlined here offers one way of addressing these challenges without sacrificing pedagogical clarity.

At the same time, the limitations identified by teachers warrant careful consideration. Difficulties in accessing ethically appropriate authentic materials, the need for additional scaffolding for lower-proficiency learners, and constraints imposed by existing curricula all point to areas for further exploration. Future work might examine how multimodal awareness develops across proficiency levels, or how shared resources and institutional support could facilitate wider use of authentic spoken interaction in classroom practice.

Overall, the findings suggest that authenticity need not remain an aspirational ideal reserved for advanced learners. With careful mediation, authentic spoken interaction can become a regular and meaningful component of ELT instruction. By foregrounding the interactional and multimodal nature of communication, the approach described here contributes to ongoing ELT discussions about how learners can be supported in making sense of real-world language use. In doing so, it offers a practical illustration of how spoken interaction can be taught not only as linguistic input, but as a socially situated practice shaped through coordinated verbal and non-verbal choices.

List of references

- Aleynikova, M. I. 2015. ‘Video films as a means of teaching a foreign language’. *Perspektivy Nauki i Obrazovaniya* 6(18): 87–90.
- Bakieva, G. and Sh. Temirova. 2020. ‘Using authentic video to develop language skills’. *Society and Innovation*: 330–4.

Candlin, C. and S. Sarangi (eds.). 2011. Handbook of Communication in Organisations and Professions. Berlin: De Gruyter.

Chuprakova, E. V. and E. S. Popovich. 2014. 'Linguistic and extralinguistic aspects of translation'. Vestnik KRU MVD Rossii 1(23): 121–4. Available at: <https://cyberleninka.ru/article/n/lingvisticheskie-i-ekstralingvisticheskie-aspekty-perevoda>

Gilmore, A. 2007. 'Authenticity in materials and practices'. ELT Journal 61/2: 97–114.

Heritage, J. and D. W. Maynard (eds.). 2006. Communication in Medical Care: Interaction between Primary Care Physicians and Patients. Cambridge: Cambridge University Press.

Dj, Israilov J., and Israilova I. Kh. "Professional medical texts as a medium for advancing lexical competence in ESP-based medical education." *Медицинский журнал молодых ученых* 15 (09) (2025): 231-234.

Исраилова, И. Х. (2025). Метафорические и фреймовые модели как средство когнитивно-дискурсивного осмысления медицинских знаний в междисциплинарной образовательной среде. *pedagogik tadqiqotlar jurnali*, 7(1), 238-243.

Исраилова, И. Х. (2021). Обогащение лексической составляющей английского языка при просмотре аутентичных сериалов по специальности. In *Діалог культур у Європейському освітньому просторі*. Київський національний університет технологій та дизайну.

Исраилова, И. Х. (2022). Требования педагогического эксперимента и его соответствие на примере исследовательской работы по обучению английского языка посредством медиа технологий. *Ilm sarchashmalari*, (7), 139-142.

Israilova, I. (2026, May). Developing specialised vocabulary in ESP through professional texts: a genre-informed, technology-supported input-processing-output model. In *Frontiers in Education* (Vol. 11, p. 1829270). Frontiers Media SA.

Ren, Z., Wu, Z. A., Wang, Y., Jakhongirkhon, I., Zhou, Q., & Du, J. (2025). Enhanced external counterpulsation intervention induces the variation of physiological parameters and shear stress metrics in the carotid artery. *Bioengineering*, 12(4), 386.

Norris, S. 2004. Analyzing Multimodal Interaction: A Methodological Framework. London: Routledge.

Syromolotova, E. M. 2014. 'Influence of extralinguistic factors on the development of modern English'. International Student Scientific Forum.

