



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

A MODEL FOR DEVELOPING STUDENTS' ORAL COMMUNICATIVE COMPETENCE IN ENGLISH LANGUAGE LEARNING ON THE BASIS OF ARTIFICIAL INTELLIGENCE PLATFORMS

Mushtariy Sultonova Farxodbek qizi

English language teacher, InterNation Lyceum

mushtariy.aliyeva@mail.ru

Abstract

The article presents a pedagogical model for the formation of oral communicative competence in students who learn English with the support of artificial intelligence platforms. The author analyses the structural components of oral communicative competence, describes the didactic functions of conversational AI systems, automatic speech recognition services and adaptive feedback tools, and proposes a four stage model of instruction that integrates these tools into the ordinary classroom cycle. The model is built on three interrelated blocks: the target and content block, the technological and procedural block, and the diagnostic and corrective block. Particular attention is paid to the distribution of roles between the teacher and the digital environment, to the criteria of speech quality assessment, and to the risks of uncontrolled use of generative systems. The pedagogical experiment carried out among lyceum students demonstrated a measurable growth of fluency, lexical range and interactive readiness in the experimental group. The results allow the author to formulate practical recommendations for teachers who intend to introduce artificial intelligence tools into the teaching of speaking.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

Keywords: Oral communicative competence, artificial intelligence, English language teaching, speaking skills, adaptive feedback, automatic speech recognition, digital learning environment, pedagogical model, formative assessment, lyceum education.

INTRODUCTION

The ability to speak English fluently and appropriately has become one of the most demanded practical outcomes of language education. Employers, universities and international certification bodies evaluate a graduate not by the volume of grammatical knowledge that the graduate is able to reproduce, but by the readiness to enter a conversation, to hold a position, to explain an idea and to react to an unexpected question. This shift of priorities has been reflected in national curricula, yet school and lyceum practice still shows a persistent gap between the declared communicative goal and the real distribution of classroom time. Observation of ordinary lessons confirms that the average learner produces only a few minutes of connected speech per lesson, and a considerable part of that speech consists of reproduced textbook phrases rather than independently constructed utterances.

The reasons for this gap are well known to every practising teacher. Large groups make individual speaking practice organisationally difficult. Learners with a low level of confidence avoid speaking in front of peers because of the fear of making mistakes in public. The teacher physically cannot listen to twenty five learners, record their errors and give each of them detailed feedback within forty five minutes. As a result, the evaluation of oral performance often remains impressionistic, and the learner receives a mark without a clear understanding of what exactly should be improved in pronunciation, grammatical accuracy or discourse organisation.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

Artificial intelligence platforms offer a realistic instrument for narrowing this gap. Modern conversational systems built on large language models are able to sustain a coherent dialogue on almost any topic, to adapt the complexity of their language to the level of the interlocutor, to reformulate a question when the learner does not understand it and to produce an analytical commentary on the learner's utterance. Automatic speech recognition services transcribe the learner's speech with a high degree of accuracy and make it possible to measure objective parameters such as speech rate, pause distribution and lexical variety. Text to speech systems provide an unlimited amount of listening input in different accents and speaking styles. Taken together, these technologies create an environment in which every learner can practise speaking individually, repeatedly and without the social pressure of the classroom.

At the same time, the simple presence of a technology in a classroom does not automatically produce a pedagogical effect. Uncontrolled use of generative systems may lead to the substitution of the learner's own speech production by a ready made text, to the formation of dependence on external prompts and to the loss of the interactive dimension of communication, which presupposes a living partner with unpredictable reactions. Therefore the central methodological task is not the selection of a particular application, but the construction of a model that defines the place of the digital tool in the instructional cycle, the functions that remain with the teacher, and the mechanisms of monitoring the real growth of the learner's speaking ability.

The aim of this article is to describe such a model. The author formulates the structure of oral communicative competence as the target of instruction, defines the didactic functions of artificial intelligence tools in relation to each component of this structure, presents the stages of the instructional cycle, and reports the results of an experimental verification of the model in the conditions of a lyceum.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

The notion of communicative competence entered language pedagogy as a reaction to the narrowly grammatical understanding of language proficiency. The classical description distinguishes grammatical, sociolinguistic, discourse and strategic components, and this four component structure remains the most workable framework for the design of speaking instruction [1]. Later interpretations added the pragmatic dimension and emphasised that competence is manifested only in performance, that is, in the observable act of communication, which makes the quality of the learner's utterance the only reliable evidence of learning [2].

Research on the teaching of speaking consistently shows that fluency develops through repeated performance of similar communicative tasks under conditions of gradually increasing pressure, while accuracy develops through focused attention to form at the moment when the learner is ready to notice the gap between the intended and the produced utterance [3]. This distinction is methodologically important for the present model, because artificial intelligence tools serve the two goals in different ways: unlimited dialogue practice serves fluency, whereas automatic analytical feedback serves accuracy and complexity. Studies devoted to computer assisted language learning have accumulated evidence that technologically mediated interaction lowers the level of speaking anxiety and increases the willingness to communicate, particularly among learners with a low self assessment of their language ability [4]. The absence of a peer audience and the possibility of repeating an utterance as many times as necessary remove the social cost of error. Investigations of automatic speech recognition in pronunciation training have demonstrated that immediate visual representation of the learner's speech, in the form of a transcript with marked



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

deviations, accelerates the correction of segmental errors more effectively than a delayed oral remark of the teacher [5].

Recent work on generative language systems in education describes both the opportunities and the risks of these tools. Authors note the high quality of the linguistic input that such systems produce, their patience and availability, and their capacity to individualise the level of difficulty [6]. At the same time researchers warn against three typical distortions: the illusion of progress, when the learner mistakes a comfortable interaction with a tolerant machine for real communicative development; the loss of authorship, when the system produces the content that the learner should have produced; and the narrowing of pragmatic experience, because a machine does not reproduce the full range of social reactions that a human interlocutor displays [7].

The methodology of blended and adaptive learning provides the organisational framework within which these opportunities and risks can be balanced. The principle of the inverted instructional cycle, in which mechanical and individual work is transferred to the digital environment while the classroom is reserved for genuine interaction, has proved effective in a number of experimental programmes [8]. Formative assessment theory supplies the second organising principle: feedback influences learning only when it is specific, timely and connected with a clearly formulated criterion that the learner is able to apply independently [9]. Finally, works on the digitalisation of education in the national context emphasise the necessity of adapting imported technological solutions to the concrete conditions of the institution, including the level of technical equipment, the language background of learners and the methodological readiness of teachers [10].



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

The analysis of these sources allows the following theoretical position to be formulated. An artificial intelligence platform is not a replacement for the teacher and not an independent method of instruction. It is a component of the learning environment that performs three functions which the teacher cannot perform in full within the limits of a lesson: the provision of an unlimited amount of individual interactive practice, the instant analytical processing of the learner's oral production, and the adaptation of the level of difficulty to the current state of the individual learner. All other functions, and above all goal setting, motivation, the organisation of genuine social interaction and the ethical evaluation of the content, remain with the teacher.

THE STRUCTURE OF ORAL COMMUNICATIVE COMPETENCE AS THE TARGET OF THE MODEL

The design of any instructional model begins with a precise description of the result that is to be achieved. In the present work oral communicative competence is understood as an integrative characteristic of the personality that includes five interrelated components.

The linguistic component covers the phonetic, lexical and grammatical resources that the learner is able to activate in real time. The essential word in this definition is activation. A learner may recognise a grammatical structure in a written test and still be unable to use it while speaking, because oral production requires automated access to the form under conditions of limited planning time.

The discourse component covers the ability to organise an utterance beyond the level of a single sentence: to open and close a conversation, to connect ideas with appropriate linking devices, to distribute information between the known and the new, and to maintain the logical line of an argument over several turns.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

The sociolinguistic and pragmatic component covers the appropriateness of speech to the situation: the choice of the register, the level of politeness, the degree of directness, and the culturally determined conventions of requesting, refusing, disagreeing and expressing gratitude.

The strategic component covers the ability to keep communication going when linguistic resources are insufficient: paraphrasing, asking for clarification, using approximations and filling pauses without breaking the flow of the conversation.

The psychological component covers the readiness to enter communication, that is, the willingness to speak, the tolerance of uncertainty and the absence of paralysing fear of error. Practice shows that in the conditions of school and lyceum education this component is often the decisive one, because a learner with sufficient linguistic resources may remain silent for purely psychological reasons. Each of the five components requires its own type of practice and its own instrument of measurement. The value of the proposed model lies in the fact that artificial intelligence tools are distributed among the components in accordance with their real capacities, rather than applied indiscriminately to the whole process of instruction.

THE STRUCTURE OF THE MODEL

The model consists of three blocks that operate simultaneously and are connected by direct and reverse information flows.

The target and content block defines what is taught. It contains the communicative goal of the cycle, formulated as an observable speech behaviour, the inventory of the linguistic means that serve this goal, and the set of communicative situations in which the goal is realised. The formulation of the goal in behavioural terms is a mandatory requirement of the model. A goal such as the study of the topic of environmental protection cannot be measured, whereas



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

a goal such as the ability to state a position on an environmental problem, to support it with two arguments and to respond to one counterargument within two minutes of connected speech defines both the content of instruction and the criterion of its success.

The technological and procedural block defines how the goal is achieved. It contains the sequence of instructional stages, the distribution of tasks between classroom work and individual work with the digital environment, and the specification of the artificial intelligence tools that serve each stage. Four categories of tools are used: conversational systems for interactive practice, speech recognition services for the transcription and objective analysis of the learner's production, text to speech systems for the generation of listening input, and analytical services for the evaluation of lexical and grammatical complexity. The diagnostic and corrective block defines how progress is measured and how deviations are corrected. It contains the criteria of the evaluation of oral performance, the instruments of the collection of evidence, the procedure of the analysis of the learner's speech portfolio and the mechanism of the individual correction of the learning route. The distinctive feature of the model is the accumulation of the learner's recorded oral production throughout the course, which converts assessment from a single examination event into a longitudinal observation of growth.

THE INSTRUCTIONAL CYCLE: FOUR STAGES

The first stage is the stage of input and modelling. The learner receives authentic or adapted listening material on the topic of the cycle, together with a transcript. The digital environment is used here for the generation of additional input adapted to the individual level: the same content is produced at different speeds and in different registers, which allows every learner to work with a text that is



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

comprehensible yet slightly above the current level. The teacher formulates the communicative goal of the cycle, demonstrates a model performance and draws the attention of the learners to the linguistic means that make this performance successful. The output of the stage is the learner's understanding of what is to be produced and by what means.

The second stage is the stage of controlled individual practice in the digital environment. The learner performs a series of short speaking tasks in interaction with a conversational system that has received a precise instruction defining the role, the topic, the level of language and the type of feedback. This instruction is prepared by the teacher and remains identical for the whole group, which guarantees the comparability of the results. The learner speaks, the system transcribes the speech, responds in the role assigned to it and, at the end of the session, produces a structured commentary: the errors that impede understanding, the errors that do not impede understanding, the lexical items that could have been used more precisely and the discourse devices that were missing. The learner repeats the task after studying the commentary. The repetition is a mandatory element: feedback that is not followed by a new attempt does not influence the formation of the skill.

The third stage is the stage of genuine classroom interaction. The tasks performed with the machine are transferred to communication with human partners in pairs, groups and whole class formats. This stage is not reducible to the previous one and cannot be delegated to the digital environment. Only a human partner produces unpredictable reactions, misunderstanding, disagreement, emotional colouring and the pressure of real time, and only in this situation does the strategic component of competence develop. The teacher observes, records the typical difficulties of the group and organises the immediate correction of those errors that are common to several learners.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

The fourth stage is the stage of assessment and reflection. The learner performs the final communicative task of the cycle, which is recorded. The record is transcribed automatically and evaluated by two instruments: the analytical criteria applied by the teacher and the objective quantitative parameters computed from the transcript. The learner compares the final record with the record made at the beginning of the cycle, formulates the personal result in words and, together with the teacher, defines the priority of the next cycle. This reflective procedure develops the learner's ability to evaluate personal speech independently, which is the basis of autonomous language development after the end of formal instruction.

CRITERIA AND INSTRUMENTS OF ASSESSMENT

The model uses a combined system of assessment in which qualitative expert judgement is supported by quantitative data extracted from the transcripts of the learner's speech.

The qualitative criteria correspond to the components of competence described above. Fluency is evaluated by the smoothness of the utterance and the absence of long unmotivated pauses. Accuracy is evaluated by the proportion of errors that impede understanding. Range is evaluated by the variety of lexical and grammatical means used in the performance of the task. Coherence is evaluated by the logical organisation of the utterance and by the use of connecting devices. Interaction is evaluated by the ability to respond to the partner, to develop the partner's idea and to repair a communication breakdown. Each criterion is described in a four level scale with verbal descriptors that are made available to the learners at the beginning of the course, because a criterion that is hidden from the learner cannot regulate the learner's behaviour.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

The quantitative parameters are computed from the transcript and include the speech rate expressed in words per minute, the mean length of run, that is, the number of words between two pauses exceeding one second, the type token ratio as an indicator of lexical variety, the proportion of subordinate clauses as an indicator of syntactic complexity, and the number of self corrections. These parameters are not used for the assignment of a mark, because their isolated interpretation may be misleading, but they provide an objective picture of the dynamics of the individual learner over the course and reveal changes that are difficult to notice through direct observation.

EXPERIMENTAL VERIFICATION

The model was verified in the conditions of a lyceum during one academic semester. Two groups of learners of a comparable initial level participated in the work. The experimental group studied according to the described model, while the control group followed the ordinary programme with the same thematic content and the same number of contact hours. The initial and final levels of oral competence were measured by an identical procedure that included a monologue task and a paired interaction task, both recorded and evaluated by two independent assessors according to the criteria described above.

The comparison of the initial and final results demonstrated a growth of the average indicators in both groups, which was expected, since both groups continued to study. However the character and the magnitude of the growth differed. In the experimental group the most significant changes were recorded in the parameters of fluency and of the willingness to communicate. The mean speech rate in the monologue task increased noticeably, the length of unfilled pauses decreased, and the number of learners who refused to perform the interaction task fell to zero, whereas at the beginning of the semester several



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

learners in both groups had declined to speak. The growth of lexical variety was also higher in the experimental group, which is explained by the systematic analytical feedback on the precision of word choice.

At the same time the growth of grammatical accuracy in the experimental group did not significantly exceed the corresponding indicator of the control group. This result corresponds to the theoretical expectation: accuracy develops through focused conscious work on form, and the automatic feedback of a conversational system, however detailed, does not replace the systematic explanation and the controlled practice organised by the teacher. This observation confirms the central methodological thesis of the model, according to which the digital environment extends the possibilities of practice but does not take over the instructional functions of the teacher.

The qualitative part of the study, based on the observation of lessons and on the interviews with the participants, revealed an additional effect that was not planned in the design of the experiment. Learners reported that individual work with the conversational system reduced the fear of speaking in the classroom, because the phrases that had already been produced several times in the digital environment were perceived as tested and safe. This observation supports the conclusion that the psychological component of competence is the one that benefits most from the introduction of artificial intelligence tools into the teaching of speaking.

RISKS AND CONDITIONS OF EFFECTIVENESS

The practical implementation of the model revealed a number of conditions without which the expected effect is not achieved.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

The first condition is the precise formulation of the instruction given to the conversational system. A vague instruction produces a vague interaction. The instruction must define the role of the system, the topic, the level of language, the length of the turns, the prohibition of switching to the native language of the learner and the exact format of the final commentary. The preparation of such instructions is a new professional skill of the teacher and requires a separate stage of methodological training.

The second condition is the prohibition of the substitution of the learner's production. The system must not produce the text that the learner is expected to produce. This limitation is achieved through the design of the task and through the format of the feedback, which indicates the direction of the improvement without supplying the improved version in a ready made form.

The third condition is the preservation of the human interaction stage. If the digital environment displaces genuine classroom communication, the strategic and pragmatic components of competence do not develop, and the learner acquires the ability to speak with a machine rather than with people. The proportion of individual digital practice in the total volume of speaking work should not exceed approximately one half of the cycle.

The fourth condition is the critical evaluation of the linguistic material produced by the system. Generative systems occasionally produce inaccurate factual statements and stylistically inappropriate formulations. The teacher remains the final authority in matters of linguistic norm and factual accuracy, and the learners should be explicitly informed about this limitation of the technology.

The fifth condition is the observance of the requirements of data protection. The recordings of the learners' speech constitute personal data, and their collection, storage and processing require the informed consent of the learners and of their



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

legal representatives as well as the observance of the internal regulations of the educational institution.

CONCLUSION

The study permits the following conclusions to be drawn.

The formation of oral communicative competence on the basis of artificial intelligence platforms is effective when the technology is integrated into a defined instructional model rather than used as an isolated supplement to the lesson. The model proposed in the article consists of the target and content block, the technological and procedural block and the diagnostic and corrective block, and it is realised through a four stage cycle that includes input and modelling, controlled individual practice in the digital environment, genuine classroom interaction and assessment with reflection.

The didactic potential of artificial intelligence tools is distributed unevenly among the components of competence. The greatest effect is observed in the psychological component, in fluency and in lexical range, that is, in those aspects of speaking that depend on the volume of individual practice and on the immediacy of feedback. The development of grammatical accuracy and of the pragmatic component continues to depend primarily on the systematic work of the teacher and on interaction with human partners.

The effectiveness of the model is conditioned by the precision of the instructions given to the digital system, by the preservation of the learner's authorship of the produced speech, by the retention of the stage of genuine human communication, by the critical evaluation of the material generated by the system and by the observance of the requirements of personal data protection.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

The practical value of the results consists in the possibility of applying the described cycle, the criteria of assessment and the set of conditions in the work of secondary schools, lyceums and language courses without substantial additional expenditure, since the model relies on widely accessible technological tools and does not require the reorganisation of the curriculum.

REFERENCES

1. Canale M., Swain M. Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing // *Applied Linguistics*. 1980. Vol. 1, No. 1. P. 1–47.
2. Bachman L.F., Palmer A.S. *Language Assessment in Practice*. Oxford: Oxford University Press, 2010. 510 p.
3. Skehan P. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press, 1998. 324 p.
4. MacIntyre P.D., Clément R., Dörnyei Z., Noels K.A. Conceptualizing Willingness to Communicate in a L2 // *The Modern Language Journal*. 1998. Vol. 82, No. 4. P. 545–562.
5. Levis J.M. Computer Technology in Teaching and Researching Pronunciation // *Annual Review of Applied Linguistics*. 2007. Vol. 27. P. 184–202.
6. Kohnke L., Moorhouse B.L., Zou D. ChatGPT for Language Teaching and Learning // *RELJ Journal*. 2023. Vol. 54, No. 2. P. 537–550.
7. Godwin-Jones R. Emerging Technologies: Partnering with Intelligent Systems for Language Learning // *Language Learning and Technology*. 2022. Vol. 26, No. 2. P. 1–19.
8. Bergmann J., Sams A. *Flipped Learning: Gateway to Student Engagement*. Washington: ISTE, 2014. 169 p.



International Conference on Education, Psychology and Humanities

Hosted Online from Moscow, Russia

Date: 28th September, 2026

Website: <https://econferencia.com>

9. Black P., Wiliam D. Assessment and Classroom Learning // Assessment in Education: Principles, Policy and Practice. 1998. Vol. 5, No. 1. P. 7–74.

10. Muslimov N.A., Usmonboyeva M.X. Innovative Educational Technologies and Pedagogical Competence. Tashkent: Fan va texnologiya, 2020. 208 p.