

AN INTERVIEW WITH THE MACHINE-INQUIRING SPEECH IN CHATGPT: THE SELF-DESCRIPTION OF ITS CAPABILITIES BY AN AI CHATBOT

[Fee-Alexandra HAASE](#)

Professor, Ph.D.

(University of Nizwa, Oman)

F.A.Haase@gmx.de, <https://orcid.org/0000-0003-4698-6955>

Abstract

The following article presents the results of an inquiry into how the chatbot ChatGPT handles speech processed in digital format. Based on ChatGPT's responses, it can be summarized that speech is treated as a human form of communication that can be analyzed using computational methods of an underlying architecture of transformer neural networks and distinct computational algorithms and models when prompted to perform a specific analytical task for spoken and written speech. In the case of an automated translation by ChatGPT we show how speech is processed. We argue that human speech-like text is generated by AI combining processes of machine learning, natural language processing, text representing speech as input, and the user's prompt. In the discussion of the results we present a distinction between human speech and AI generated speech using the concept of fabrication for describing the generated content. In a model we illustrate the distinct areas of human speech and digital information.

Keywords: speech, artificial intelligence, chatbot, ChatGPT, analytics

Rezumat

Articolul prezintă rezultatele unei investigații privind modul în care chatbot-ul ChatGPT gestionează vorbirea procesată în format digital. În baza răspunsurilor prezentate de ChatGPT, putem rezuma informațiile furnizate: ChatGPT consideră vorbirea ca fiind o formă umană de comunicare pe care o poate analiza folosind metode computaționale ale unei arhitecturi subiacente de rețele neuronale de transformare, algoritmi și modele computaționale distincte atunci când i se solicită să efectueze o sarcină analitică specifică pentru vorbirea orală și scrisă. În cazul unei traduceri automate realizate de ChatGPT, arătăm cum este gestionată vorbirea. Susținem că textul asemănător vorbirii umane este generat de inteligența artificială, combinând procese de învățare automată, procesare a limbajului natural, text care reprezintă vorbirea ca input și promptul utilizatorului. Discutând rezultatele, prezentăm o distincție între vorbirea umană și cea generată de inteligența artificială, utilizând conceptul de fabricație. Într-un model, ilustrăm domeniile distincte ale vorbirii umane și ale informației digitale.

Cuvinte-cheie: vorbire, inteligență artificială, chatbot, ChatGPT, analitică

1. Introduction

1.1. Inquiring Features and Functions of Speech in the Digital Space: The Chat as an Example of the Impact of Human-Machine-Interaction

A broad area of interest regarding the contemporary practice of communication emerges from the interaction of humans and machines capable of processing and manipulating human speech in automated digital processes in a recorded,

mediated, encoded, and symbolically represented format. The digital encoding of human speech is the key activity for the following computational processes including automated computational processes performed by *ChatGPT*.

The speech style imitating human speech and used by automated AI bots is one layer of the produced output. Beneath this, automated imitative conversational speech-style processes confront data and information derived from human speech with trained models. The word *chat* is an example of the complexity of influences that form the contemporary spoken language and of the impact of human-machine interaction on speakers of English. User-generated prompts entered into ChatGPT as typed natural-language inputs are answered by ChatGPT in an imitative framework of human dialogical communication addressing the user as a conversational partner. The description of digital conversational communication as *chat* was first employed to describe conversations between humans in digital applications, e.g., social media. Chat as a type of conversation in the English language can be traced back to the English language before the emergence of digital communication. The use of the word *chat* is an example of the semantic range across the areas of natural language use, cutting across a range of applications with meanings related to spoken and recorded human speech in a familiar way, commonly associated with a negative connotation.

In the *Collins English Dictionary* (2012) the use of *chat* in American and British English is described. *Chat* as a verb in American English has the meaning 'converse in a familiar or informal manner'. In digital technology the meaning is "to participate with one or more people, through the internet, in a real-time conversation, typically as a series of short text exchanges in a specific application, as instant messaging, or by using images, voice, video, or some combination of these". As a noun the meaning of *chat* in American English is 'an informal conversation'. In digital technology it depicts "a real-time conversation between two or more people over the internet or another computer network". The verb *chat* in British English is used with the meaning 'talk in an easy familiar way' and 'exchange messages in a chatroom'. The noun *chat* in British English is used for 'an informal conversation or talk conducted in an easy familiar manner', 'the exchange of messages in an internet or other network chatroom', 'any Old World songbird of the subfamily Turdinae (thrushes, etc.) having a harsh chattering cry', and 'any of various North American warblers, such as *Icteria virens* (yellow-breasted chat)' as well as 'any of various Australian wrens (family Muscicapidae) of the genus *Ephthianura* and other genera'.

In the *Oxford English Dictionary* (n.d.) the verb *chat* is traced to its earliest occurrence in the *York Plays* (1450). Until the 17th century *chat* had the meanings 'engage in trivial or frivolous talk' and 'prattle'. An onomatopoetic origin of *chat* can be supported by the assumption that it is an abbreviation of the word *chatter* representing the sound of birds as well as human speech in a

familiar tone. The *Collins English Dictionary* (2012) discusses the American and British English variants of the verb *chatter*: In American English this verb without an object has the meanings 'talk rapidly in a foolish or purposeless way', 'jabber', and 'utter a succession of quick, inarticulate, speech-like sounds, as monkeys or certain birds'. As a verb used with an object *chatter* has the meanings 'utter rapidly or purposelessly' and 'cause to chatter, as the teeth from cold'. The noun *chatter* has the following meanings: 'purposeless or foolish talk', 'a series of waves or ridges on the surface of a piece of metal that has been imperfectly drawn or extruded', 'the act or sound of chattering', and 'online, phone, radio, or other electronic communication among people, often involving a harmful political activity such as espionage or terrorism'. The verb *chatter* in British English has the meanings 'speak (about unimportant matters) rapidly and incessantly', 'prattle', '(intr) (of birds, monkeys, etc) to make rapid repetitive high-pitched noises resembling human speech', '(intr) (of the teeth) to click together rapidly through cold or fear', and '(intr) to make rapid intermittent contact with a component, as in machining, causing irregular cutting'. In the 1990s the word *chat room* came to be used for the virtual places where online real-time conversations took place. An example for the impact of quasi-technical terms describing communicative exchange is the use of *chat*. Originally describing an informal conversational talk, the word experienced a revival as a term for online conversations among humans and later developed an extended meaning among young people, who may use *chat* to refer to a person or group of people. As M. Gordoni mentions in her article *The Slang Evolution: Why Gen Alpha Calls Each Other 'Chat'* in *Movieguide* (June 30, 2025), the use of the word *chat* among young people derives from the terminology of the World Wide Web. "Chat" refers to an audience of people. It originated from YouTube streamers addressing their audience, with expressions such as "What's up, Chat?" as reported by Parade. Streamers' viewers talk in the comment section, also known as the chat. Gradually, streamers began to address the chat almost as if it were a single person. Originally associated with a negatively connotated form of talk among humans in an informal register, in the sociolect of young people *chat* has become a word used to refer to people participating as a group in an online chat, thereby reducing individuals to their communicative function as participants in a digital interaction.

1.2. Researching Inquiring Speech Acts in Human-Machine Environments: How Does ChatGPT Handle Speech, Prompts and Answers?

The following research outcomes describe the current assessment of AI applications as means of analyzing speech. The tendency is that AI applications are considered to be means of compensating for missing aspects in human speech performance, whether at the level of learning public speaking with AI applications, or as a part of a general perspective in which AI is

viewed as a technology that is permanently improving its ability to simulate human features and to blur the boundaries between human and machine. As concrete applications of AI related to human speech the industry has produced in 2025 countless apps that aim at improving human public speaking through automated AI. As a reflection of this trend in research, D. Tiwari, P. P. Arora, Sh. Geetey, L. Shrivastava, K. K. I. Shree, D. L. Melon, and S. R. Nayagi in a study on public speaking instruction with AI applications state that “in the era of accelerated technological advancement, artificial intelligence (AI) has become a significant enabler of transformation in human communication, with notable applications in public speaking training” (Tiwari *et alii*, 2025, p. 1401). These researchers list applications of AI: “Examples include virtual coaching systems, speech rehearsal simulators, and natural language processing (NLP) frameworks that assess communicative clarity, rhetorical effectiveness, non-verbal delivery, and persuasive potential”. Z. Zhang in *AI-Powered Intelligent Speech Processing: Evolution, Applications and Future Directions* describes in a diachronic overview of the development of AI the progression of the adaptation of human speech features by AI applications (Zhang, 2025, p. 918). Z. Zhang here compiles “an overview of the historical evolution of speech recognition, synthesis, and processing technologies, highlighting the transition from statistical models to deep learning-based models” from “rule-based and statistical models of the 1960s” to “deep learning models, such as deep neural networks (DNN), convolutional neural networks (CNN), and recurrent neural networks (RNN)”. The researcher claims that “these advancements have led to more natural and accurate speech outputs” and mentions as application domains of AI-driven speech processing “smart homes, intelligent transportation, healthcare, and finance”. “Integrating AI with technologies like the Internet of Things (IoT) and big data”, speech technology “is being applied in voice assistants, autonomous vehicles, and speech-controlled devices”. As key challenges Z. Zhang mentions “the scarcity of data for low-resource languages” and “concerns related to data privacy”, “algorithmic bias”, and “legal responsibility”. D. Inserra in the *Briefing Paper No. 178* (July 16, 2024), *Artificial Intelligence Regulation Threatens Free Expression* of the *Cato Institute* stated that “with AI set to impact so much of our lives, the decisions over what kinds of speech AI should or should not be allowed to generate will likely become far greater battles than those fought over social media content policies”. The paper reflects that “as exemplified by ChatGPT, new generative AI technologies have gained widespread public usage in various domains, including productivity, art, education, commerce, and science”. Here the implicit understanding is expressed that AI is capable of generating speech. Research approaches that look at the impact of AI applications on human language acquisition also exist. So H. Rafique, I. Nazeer, and J. Rehman in *Impact of ChatGPT on Lan-*

guage Evolution: A Linguistic Analysis took the approach of aiming “to identify changes in terms of word choice, sentence structure, and semantics introduced by ChatGPT and how the functionality of ChatGPT impacts the pragmatics of the English language” (Rafique *et alii*, 2024, p. 56). As a result it is stated that “ChatGPT generates neologisms, unique cultural references, verbal contextual adaptations, and a wide range of technical terms” and “that user preferences for communication mediums vary, which indicates a complex field of users and their expectations in the context of language generated by ChatGPT”, while “it was observed that ChatGPT contributes significantly to developing the pragmatic aspects of language use in the modern era”.

These examples of research on AI reflect the overall state of the art in AI applications for speech. We use ChatGPT in our study concerning the relation of contemporary AI and speech. Our approach examines the distinction between speech as human communication and what AI generates while carrying features of speech, focusing on the differences between the two types of communication. Since we inquire the self-understanding of this application concerning its methods and approaches to speech for analytical tasks in a digital format and the boundaries between speech and what is no longer speech, we address here our working definition of speech as a realized and manifested capability of the human body to communicate with spoken or written expressions stemming from this capacity. On the contrary, the imitation in a conversational style via the interface exists between a human sending prompts and answers delivered by ChatGPT in a default framework of human-machine communication. In the case of ChatGPT it is a text construct that does not qualify as speech according to our definition above. The questions asked by prompts on December 17, 2025 are:

- A. How does ChatGPT define speech?
- B. Which computational methods and algorithms does ChatGPT use in order to analyze a digital text submitted to ChatGPT via a prompt requesting analysis of its features as written speech?
- C. How does ChatGPT assess and classify a unit of spoken or written speech that is digitally presented to it?
- D. Which linguistic methods can ChatGPT use to analyze spoken or written speech that is digitally presented?
- E. What can ChatGPT do to analyze a digital text?

We will present and assess the results from the prompted questions of ChatGPT in a contrastive study. The results are compared to the methods of language studies at the level of terminology used and discussed in a case study of a translation generated by ChatGPT that evaluates the concept of speech in this context. In order to distinguish between speech and speech-like features in automatically generated text by this AI application we use

the concept of fabrication and a model to illustrate both boundaries and interactions between these modes of communication.

2. Speech Analysis in ChatGPT: Prompts and Responses

2.1. Self-Description of ChatGPT Concerning its Definition of Speech and its Computational Methods and Means

A. Self-Description of ChatGPT Concerning its Definition of Speech

The concept of speech is for ChatGPT a traditional one, understood as a process of human communication. To the question "How does ChatGPT define speech?" (Question A) ChatGPT provides definitions in the following responses (December 17 and 21, 2025):

"Speech is typically defined as a form of human communication that uses spoken language—sounds produced by the vocal tract—to express thoughts, ideas, emotions, or information" [...].

A concise definition would be: "Speech is the vocal expression of language used by humans to communicate meaning" (December 17, 2025).

In a general, interdisciplinary sense, ChatGPT defines speech as *the intentional, structured use of spoken language sounds to produce meaning in communicative interaction*.

This definition can be expanded across several theoretical perspectives that are often relevant in linguistics, rhetoric, and communication studies.

Concise working definition: "Speech is the situated, intentional, and temporally organized use of spoken signs to enact meaning and communicative effects between participants" (December 21, 2025).

In both versions no references are attached to the results retrieved from the prompt. Structurally, both answers to the prompt show a similar sequence of sections in the response: a general definition, a breakdown of areas of application of speech in science, a *concise definition* (December 17, 2025) and a *concise working definition* (December 21, 2025). While the definitions and breakdown in the middle part of the first version entail a focus on linguistics and emphasize that speech is a human feature, the definitions and breakdown in the middle part of the second version mention that speech is the use of "spoken signs". While the first version refers to speech as a linguistic phenomenon, the second version breaks speech down from different perspectives, including a linguistic perspective as the "oral realization of language (as opposed to writing)", a communication-theoretical perspective with speech as "a semiotic act that encodes meaning through conventional signs", a rhetorical perspective as *oratio, logos, a pragmatic and philosophical perspective with speech as "form of action (speech act)"*, and a semiotic perspective, in which "speech can be defined as [...] a temporally unfolding sign complex in which acoustic signs are organized to generate meaning, tension, and effects within a communicative system". The results compile common terms

of the specific branches of sciences in a general way without applying a clear scientific methodology, presenting them as general statements without any examples, references or the level of analytical treatment expected from a scientific definition in a reference work or from a working definition in a scientific study.

B. Self-Description of ChatGPT Concerning its Computational Methods and Means

Upon the question "Which computational methods and algorithms does ChatGPT use in order to analyze a digital text that was submitted to ChatGPT with a prompt to *analyze* its features as written speech?" (Question B), ChatGPT answers (December 17, 2025):

Here's a clear, technical but non-proprietary explanation of the computational methods and algorithms that ChatGPT (including when it analyzes a submitted digital text to examine features of written speech) uses in broad terms.

In the following sections ChatGPT lists six elements:

1. Underlying Architecture – Transformer Neural Networks
2. Pre-training and Fine-tuning
3. Text Representation – Embeddings
4. Attention Mechanisms
5. Task Processing Flow
6. Algorithms and Techniques Involved

In point 1 *Underlying Architecture – Transformer Neural Networks* ChatGPT describes that "transformers process text by relating each token (word or piece) to every other token using self-attention mechanisms". In point 2 *Pre-training and Fine-tuning* ChatGPT states that the model is trained on very large corpora using self-supervised learning" with the objective to "predict missing or next tokens based on context", including aspects of grammar and syntax, semantic associations, general world knowledge, discourse patterns, and fine-tuning. In point 3, *Text Representation – Embeddings*, the text of the response explains that when a text is submitted, "each token (word or sub-word) is converted into a vector embedding" and these vectors encode lexical meaning, contextual role, and position in sequence. ChatGPT describes that "these representations are the basis for all downstream analysis". In point 4, *Attention Mechanisms*, the text by ChatGPT states that the "self-attention mechanism lets the model weigh the importance of each token relative to others, capture long-range dependencies (e.g., references, parallel clauses, discourse patterns), and identify features such as key themes, rhetorical structure, focus shifts, and stylistic markers adding that "this is crucial when analyzing features of written speech like voice, register, emphasis, and coherence." In point 5, *Task Processing Flow*, ChatGPT breaks down the process stating that "when you ask ChatGPT to analyze text features as written

speech, it performs several internal steps: *Tokenization* that “splits text into tokens” and “converts them into numerical representations”; *Contextual Encoding* that “applies transformer layers to embed meaning and relational features”; *Pattern Recognition* that “identifies linguistic features such as “syntax (structure), semantics (meaning), tone/style, pragmatic features (speech acts, politeness, intents), discourse coherence and cohesion; *Task-Specific Reasoning* that “applies learned patterns to the analysis prompt: e.g., “Identify spoken speech features in writing” → identifies markers like cadence, repetition, fillers, register, etc.”, and *Text Generation* that “produces a response in natural language summarizing those features. In point 6, *Algorithms and Techniques Involved*, ChatGPT lists the algorithmic processes stating that “while not a full implementation list (for proprietary reasons), the key algorithmic components include: *Attention Computation, Feedforward Networks, Positional Encoding, and Layer Normalization & Residual Connections*”. As a summary (short) ChatGPT answers the prompted question as follows. When ChatGPT analyzes a text for features of written speech it does this by:

- Tokenizing the text into vectors.
- Running it through a transformer neural network with self-attention.
- Using what it has learned about language patterns during training.
- Identifying relevant features (tone, structure, style) based on contextual patterns.
- Generating a human-readable analysis using learned text generation capabilities.

The summary adds “If you want, I can also explain how one could implement a simplified version of such an analysis pipeline using open-source tools (e.g., Hugging Face Transformers, spaCy, NLTK)”. In this response ChatGPT generates text concerning proprietary constraints, stating in point 6, *Algorithms and Techniques Involved*, that not all information about them is provided. A part of the terms employed are technical (e.g., *tokenization, contextual encoding*), while others are framed using terminology associated with human cognitive processes or natural features of living and conscious beings (e.g., *text generation, task-specific reasoning, attention mechanisms*).

2.2. The Analytical Methods for Speech as Digital Text Used by ChatGPT

A. Self-Description of ChatGPT Concerning its Assessment and Classification of Spoken or Written Speech

After prompting the question “How does ChatGPT assess and classify a unit of spoken or written speech that is digitally presented to it?” (Question C), ChatGPT presents this text (Dec. 17, 2025):

ChatGPT doesn’t “hear” or directly perceive speech in the human sense. It assesses and classifies language only after it has been digitized into text or tokens. Here’s how that works, step by step, at a conceptual level.

ChatGPT breaks down these steps into seven points:

1. Input Representation (What ChatGPT Actually Receives)
2. Tokenization (Breaking Language into Units)
3. Embedding (Turning Tokens into Meaningful Numbers)
4. Contextual Analysis (Understanding the Unit of Speech)
5. Classification (Implicit, Not Label-Based)
6. Response Generation (Prediction, Not Retrieval)
7. What ChatGPT Doesn't Do

In point 1, *Input Representation (What ChatGPT Actually Receives)*, the text of the prompt-response states that if the input is spoken, “it is first converted into text using a speech-to-text system (e.g., ASR)” and if the input is written, “ChatGPT receives the text directly”. “In both cases, the input is just a sequence of characters”. In point 2, *Tokenization (Breaking Language into Units)*, the text of the prompt-response states that the “text is converted into tokens, which are” whole words (“language”), subwords (“lang” + “uage”), punctuation or symbols, and sometimes parts of words, e.g. “How does ChatGPT assess speech?” might become tokens like [“How”, “ does”, “ Chat”, “GPT”, “ assess”, “speech”, “?”]. The tokens are then “mapped to numerical IDs”. In point 3, *Embedding (Turning Tokens into Meaningful Numbers)*, the text of the prompt-response states that each token ID is “converted into a vector (a list of numbers) in a high-dimensional space”. These embeddings encode semantic meaning, grammatical role, typical usage context, and relationships to other tokens. In point 4, *Contextual Analysis (Understanding the Unit of Speech)*, the text of the prompt-response states that a transformer architecture is used. ChatGPT looks at all tokens at once, weighs how much each token relates to every other token (self-attention), and builds a contextual representation of the entire input. Here ChatGPT states that it “identifies” features such as *sentence boundaries*, *intent* (question, command, explanation, etc.), *topic*, *tone* (formal, casual, technical), *discourse role* (definition, argument, example, etc.). In point 5, *Classification (Implicit, Not Label-Based)*, the text of the prompt-response states that any classification of speech is “implicit” and “not label-based”: ChatGPT does not explicitly label input as “This is a noun phrase” or “This is persuasive speech”, but classification is implicit: “Patterns learned during training guide how it responds” and “The model “recognizes” types of units because similar token patterns led to certain outputs during training”. Examples of implicit classification are: a sentence ending with “?” “→ likely a question”, “Explain how...” “→ instructional/explanatory intent” and academic tone + abstract nouns “→ analytical discourse”. In point 6, *Response Generation (Prediction, Not Retrieval)*, the text of the prompt-response states that ChatGPT predicts the most likely next token given the entire context, repeats this process token by token, produces a response consistent with the inferred intent, the discourse type, and the topic and style. ChatGPT adds here that “No rules or stored definitions are consulted – just probability

over tokens.” In point 7, *What ChatGPT Doesn’t Do*, ChatGPT declares that it does “not understand meaning consciously”, “not store or recall your input as a memory”, “not apply symbolic grammar rules explicitly”, and “not “classify” speech using fixed categories unless asked to”. In a section entitled *In One Sentence*, a summary ends the text: “ChatGPT assesses a digitally presented unit of speech by converting it into tokens, embedding those tokens into numerical representations, analyzing their relationships using attention mechanisms, and implicitly classifying the input through learned patterns to generate an appropriate response”.

As for the breaking down of the process by which ChatGPT handles text, it states that it uses only probabilities associated with tokens and that no rules or definitions of speech are explicitly applied. Thus, any reference to a previously archived assessment of speech does not exist, neither explicitly as scholarly knowledge nor as a result of previously prompted and analyzed speech presented to ChatGPT. On the one hand, rejecting labels and classification, ChatGPT presents its analysis as independent of previously established knowledge, scientific methods, and from its own previous results. On the other hand, ChatGPT employs, in the description of its core functionality, its own terminological concepts, established computational terminology, and scientific terms. These terms are obviously taken for granted and used in its assessment of speech, including concepts such as discourse, topic, and tone.

B. Self-Description of ChatGPT Concerning its Linguistic Methods Employed to Analyze Spoken or Written Speech

Upon the prompt, "Which linguistic methods can ChatGPT use to analyze spoken or written speech that is digitally present?" (Question D) (December 17, 2025), ChatGPT delivers the following text:

ChatGPT can analyze digitally present spoken language (after transcription) and written language using a range of linguistic methods drawn from linguistics, NLP, and discourse analysis. Below is a structured overview, from surface-level analysis to deeper interpretive methods.

ChatGPT breaks the overview down into the following categories:

1. Phonological & Prosodic Analysis
2. Morphological Analysis
3. Lexical Analysis
4. Syntactic Analysis
5. Semantic Analysis
6. Pragmatic Analysis
7. Discourse & Conversation Analysis
8. Sociolinguistic Analysis
9. Computational/Quantitative Methods
10. Critical & Rhetorical Analysis

According to the text of the prompt-response by ChatGPT (point 1, *Phonological & Prosodic Analysis*), it “applies to spoken language after transcription, sometimes with annotations. It entails the analysis of phonemic patterns (via IPA-style or simplified transcription) and prosodic features (when marked): stress, intonation, rhythm, pauses, hesitations, elongations (e.g., “uh”, “...”), disfluency analysis, fillers, repetitions, repairs, and speech fluency vs. hesitation patterns. ChatGPT adds that “direct acoustic analysis (pitch, waveform) is not possible without audio-processing tools, but symbolic representations can be analyzed”. In the prompt-response of point 2, *Morphological Analysis*, at the “word-level structure”, ChatGPT performs analyses of morpheme segmentation (roots, prefixes, suffixes), inflectional vs. derivational morphology, word formation processes, compounding, blending, neologisms, and grammatical marking (tense, number, case, gender). In the prompt-response of point 3, *Lexical Analysis*, the *lexical analysis* refers to vocabulary richness/diversity, register analysis (formal, informal, technical, slang), semantic fields (clusters of related meanings), keyword and frequency analysis, and polysemy and ambiguity detection. In the prompt-response of point 4, *Syntactic Analysis*, the *syntactic analysis* refers to constituent structure (phrases and clauses), dependency relations, sentence complexity, coordination vs. subordination, embedded clauses, grammaticality and variation, and word order patterns (e.g., SVO, topicalization). In the prompt-response of point 5, *Semantic Analysis*, the *semantic analysis* refers to literal meaning vs. implied meaning, thematic role assignment (agent, patient, etc.), semantic coherence and cohesion, ambiguity resolution, figurative language, metaphor, metonymy, and idioms. In the prompt-response of point 6, *Pragmatic Analysis*, the *pragmatic analysis* refers to “meaning in context” and includes speech act analysis, assertions, questions, commands, promises, implicature (what is suggested but not stated), politeness strategies, deixis (this/that, here/there, now/then), presuppositions, and contextual appropriateness. In the responding text of point 7, *Discourse & Conversation Analysis*, the *discourse & conversation analysis* refers to turn-taking patterns, topic introduction, maintenance, and shift, coherence relations (cause, contrast, elaboration), anaphora and reference tracking, narrative structure, and rhetorical structure (RST). In the prompt-response of point 8, *Sociolinguistic Analysis*, the *sociolinguistic analysis* refers to register and style shifting, dialectal features (when represented in text), identity markers (professional, generational, cultural), power dynamics and stance, and code-switching (if multilingual). In the prompt-response of point 9, *Computational/Quantitative Methods*, the *computational/quantitative methods* refer to sentiment analysis, emotion classification, stance detection, topic modeling, readability metrics, and authorship attribution (probabilistic, not definitive). In the prompt-response of point 10, *Critical & Rhetorical Analysis*, the *critical & rhetorical analysis* refers to framing and

persuasion strategies, bias detection, ideological positioning, argument structure, and fallacy identification. As *key limitations*, the response of ChatGPT adds:

- No direct access to audio signals unless they are transcribed
- No ability to infer speaker intent or truth with certainty
- Cultural and contextual assumptions depend on the available data
- Analysis is probabilistic rather than deterministic

The scientific and systemic areas are mixed up by ChatGPT. For example, rhetorical elements occur under point 5, *Semantic Analysis*, while the section lacks semantic approaches stemming from the field of semantics. On the contrary, under point 10, *Critical & Rhetorical Analysis*, ChatGPT refers only to framing and persuasion strategies, bias detection, ideological positioning, argument structure, and fallacy identification, while lacking broader rhetorical approaches to speech and the rhetorical concept of speech itself as an approach to rhetorical speech analysis. We can state that the approach of ChatGPT to speech lacks general systemic approaches from the scientific disciplines relevant to the analysis of speech, while it uses random terms from these sciences, often without the systemic context in which they are generally employed.

C. Self-Description of ChatGPT concerning the Question What it Can Do to Analyze a Digital Text

As a response to the prompt "What can ChatGPT do to analyze a digital text?" (Question E) (December 17, 2025), ChatGPT answered:

ChatGPT can analyze a digital text in many ways, depending on what you want to learn from it. Here are the main types of analysis it can perform, with concrete examples:

ChatGPT lists these methods:

1. Summarization & Structure
2. Language & Style Analysis
3. Semantic & Meaning Analysis
4. Critical & Argument Analysis
5. Thematic & Literary Analysis
6. Data Extraction & Transformation
7. Sentiment & Bias Detection
8. Editing & Improvement
9. Cross-Text & Comparative Analysis
10. Custom Analytical Tasks

In point 1, *Summarization & Structure*, as stated in the text of the prompt-response, the tasks ChatGPT can perform include creating short or detailed summaries, identifying main ideas, themes, or arguments, outlining the logical structure (sections, claims, conclusions), and extracting key quotes or

passages. In point 2, *Language & Style Analysis*, as stated in the text of the prompt-response, the tasks ChatGPT can perform include analyzing tone (formal, persuasive, emotional, neutral), identifying writing style (academic, narrative, journalistic), assessing clarity, readability, and complexity, and detecting rhetorical devices (metaphors, repetition, emphasis). In point 3, *Semantic & Meaning Analysis*, as stated in the text of the prompt-response, the tasks ChatGPT can perform include explaining implicit meanings or subtext, clarifying ambiguous sentences, identifying assumptions, implications, or presuppositions, and comparing how different word choices affect meaning. In point 4, *Critical & Argument Analysis*, as stated in the text of the prompt-response, the *critical & argument analysis* is related to tasks such as identifying claims, evidence, and reasoning, spotting logical fallacies or weak arguments, evaluating coherence and consistency, and distinguishing facts from opinions. In point 5, *Thematic & Literary Analysis*, as stated in the text of the prompt-response, *thematic & literary analysis* is related to tasks such as identifying themes, motifs, symbols, analyzing characters, narrative voice, or perspective, comparing texts or authors, and interpreting literary techniques. In point 6, *Data Extraction & Transformation*, as stated in the text of the prompt-response, the *data extraction & transformation* is related to the tasks ChatGPT can perform including extracting names, dates, places, and definitions; transforming text into tables, timelines, or diagrams; classifying content by topic or category, and converting text into Q&A, flashcards, or summaries. In point 7, *Sentiment & Bias Detection*, as stated in the text of the prompt-response, *sentiment & bias detection* are related to tasks such as detecting sentiment (positive, negative, neutral, mixed), identify emotional intensity, analyzing ideological bias or framing, and comparing tone across multiple texts. In point 8, *Editing & Improvement*, as stated in the text of the prompt-response, *editing & improvement* are related to tasks such as suggesting rewrites for clarity or tone, simplifying complex passages, improve grammar, flow, or cohesion, and adapting text for different audiences (academic, general, children). In point 9, *Cross-text & Comparative Analysis*, as stated in the text of the prompt-response, *cross-text & comparative analysis* are related to tasks such as comparing multiple documents, identifying similarities, differences, and contradictions, tracking how an idea evolves across texts, and aligning a text with external frameworks or criteria. In point 10, *Custom Analytical Tasks*, as stated in the text of the prompt-response, *custom analytical tasks* are related to the interactive possibility that the user may define a framework for analysis. ChatGPT states that “you can define your own framework, such as ” using Bloom’s Taxonomy, applying discourse analysis, narrative analysis, or content analysis, scoring text against a rubric or checklist, and annotating it line by line with comments. At the end of the prompt-response under “*What ChatGPT Needs From You*”, the text states that in order “to analyze a digital text effectively, you can provide”:

The text itself (or a section of it)

Your goal (summary, critique, thematic analysis, etc.)

Any constraints (length, academic level, format)

If you'd like, you can paste a text and tell me how you want it analyzed, and I'll do it step by step.

3. Case Study, Discussion, and Model of AI and Speech

3.1. Introduction: Limits of the Analysis of Speech by ChatGPT

ChatGPT is a prominent application representing the state-of-the-art in artificial intelligence technology, including natural language processing (NLP) and large language models. The self-description of its capabilities relies to a high degree on terminology from science, while claiming self-reliance in its computational methods and independence from other knowledge derived from scientific disciplines or from knowledge generated through previous ChatGPT tasks when engaged in the analysis of a particular text representing speech. Assessing the quality of the analytical processes of speech by ChatGPT in comparison with linguistic methods of research on speech, what ChatGPT offers mimics and borrows from established research methods of different disciplines, whereas it often lacks the established distinctions between them. The criterion that a research task can be replicated and thereby attain scientific relevance is not fulfilled in ChatGPT. Additionally, in ChatGPT the sources used (the material for the research project) are not treated as documents and the methods employed are neither visible nor transparent to the user. The language models constitute besides the text of the researcher an integral element of the analysis. They cannot be called methodological elements, but they constitute a part of the research corpus, that is compared with another particular part of the research corpus, in our case the chosen text representing speech. Additionally, the AI application relies on probability and any output expresses likelihood based on a limited body of digital data. Constituted as such an obscure framework, the results from a chatbot like ChatGPT are not comparable to research performed by scientific disciplines.

A problematic stance of research applied to digital text corpora with AI applications like ChatGPT results from the individual training of large language models for each AI application. Different results are generated based on the underlying language models and computational technology of a given AI application. Further, an AI application like ChatGPT is constantly updated. The language models are continuously trained which may generate different results at a time/date A than at a time/date B. For example, a translation of a written text from L1 to L2 is subject to the linguistic range offered by the AI application. This range determines linguistic features that challenge the AI:

- a) Language pairs available
- b) The Choice of equivalent words in L1 for expressions in L2
- c) The Selection of words from the available lexical pool

- d) The Application of suitable grammar and language variety (e.g., L2 as variety of British English or American English, the lack of 'Understanding' of an older state of a language like Middle English)
- e) Syntax and style as variable linguistic features of L1 and L2
- f) The Availability of the pragmatic context of the text in L1 and L2
- g) Specific cultural meanings and their translatability between L1 and L2
- h) Access to linguistic registers by the AI (e.g., ~~specific~~ ~~specific~~ specialized terminologies)

As for specific analytic processes, the AI may perform them and deliver results (e.g., a word list or a glossary with an explanation or a translation of a text), but the resulting output lacks explicit sources or references. Automated processes provide no insight into the procedures applied after prompting, the sources used, the automated computational processes, and the language models used for text processing. In its self-description of its *ad hoc* approach to the analysis of speech, provided in response to the prompt "How does ChatGPT assess and classify a unit of spoken or written speech that is digitally presented to it?" (under point 7 *What ChatGPT Doesn't Do*) ChatGPT stated that it does "not understand meaning consciously", "not store or recall your input", and "not "classify" speech using fixed categories unless asked to". Thus, this random approach does not allow for a stable and replicable research design and results. Previous knowledge is not consulted, while ChatGPT borrows and uses scientific terminology and the results of its operations are generated from patterns learned by large language models. The output is produced in the form of a response text after a prompt and a digitally represented text sample have been provided to the AI.

3.2. A Case Study of the ChatGPT Application: Automated Translation and Glossing of 2 Peter 2:3

The performance of an automated translation by ChatGPT can be shown and assessed by comparison with a traditional translation task. The Greek text of 2 Peter 2:3 (Westcott and Hort, 1881) contains the expression *πλαστοῖς λόγοις*. Below, on the left side, the English translation of the *New International Version* of the *New Testament* of 2 Peter 2:3 (*Bible Hub*) is displayed, while on the right side, the response generated by ChatGPT to the the prompt issued on December 17, 2025 is presented:

In their greed these teachers will exploit you with fabricated stories. Their condemnation has long been hanging over them, and their destruction has not been sleeping.

Here is a clear, fairly literal English translation of the Greek text:

"And in greed they will exploit you with fabricated words; their judgment from long ago is not idle, and their destruction is not asleep".

(*New International Version*)

If you would like, I can also provide:

a more wooden/literal translation,
a more polished idiomatic translation,
or a comparison with major English
Bible translations (e.g., ESV, NASB, NRSV).

ChatGPT Prompt “Make a translation
of this text” (December 17, 2025)

2 *Peter* 2:3 is translated in the *New International Version* of the *New Testament* using the expressions “in their greed”, “these teachers will exploit you with fabricated stories” in the first sentence, and “Their condemnation has long been hanging over them, and their destruction has not been sleeping” in the second sentence. The original concept of λόγος in the Greek expression πλαστοῖς λόγοις is here translated as *stories*, a rendering that carries a connotation of vividness and introduces the notion of a narrative mode of communication by humans depicted as teachers within the translation. As for the translation of 2 *Peter* 2:3 presented by ChatGPT after being prompted (December 17, 2025), ChatGPT described it as a “clear, fairly literal English translation of the Greek text”. The expression *fabricated words* is used in the first of two sentences separated by a semicolon in the sentence “And in greed they will exploit you with fabricated words” followed by a second sentence “their judgment from long ago is not idle, and their destruction is not asleep”. The original concept of λόγος is here translated as *words* in the expression *fabricated words* for the Greek original πλαστοῖς λόγοις. ChatGPT was prompted on December 17, 2025 with the request “Make a word list of the Greek words with an English translation of each word of the following text: καὶ ἐν πλεονεξίᾳ πλαστοῖς λόγοις ὑμᾶς ἐμπορεύσονται οἷς τὸ κρίμα ἔκπαισι οὐκ ἄργεῖ, καὶ ἡ ἀπώλεια αὐτῶν οὐ νυστάζει”. ChatGPT responded (December 17, 2025) with a “word-by-word-list of the Greek text, giving a basic English translation/gloss for each word as it appears in the sentence (without providing a full grammatical analysis). Here πλαστοῖς is translated as ‘fabricated’ or ‘false’ and λόγοις as ‘words’ or ‘messages’. In this automated translation and glossing process, the results in L2 lack explicit references to a source, a commentary, or a structural organization of the lemmata.

3.3. The Concept of Fabrication: A Description for AI Generation and Resulting Digital Constructs

The Latin word *fabricatio* used to describe a physical process of making or construction, is related to the verb *fabricari* and *fabrica* meaning *workshop* and *craft*. These Latin words have reflexes in the English language in the words

fabric (textile) and *forge*. The automated processes used by AI applications produce, through a process of construction, an alternative to the original input data in the form of output data. Just as a block of wood can be cut, carved, newly assembled, or colored AI can apply different automated processes that alter an existing digital data set and transform it into newly fabricated output.

In *A Latin Dictionary* edited by Lewis and Short (1879), the word *fabricatio* has the meanings ‘a making’, ‘framing’, ‘structure’, ‘manner of construction’ with the comment “(rare but class.)”. The term occurs with its meaning of a physical construction in Vitruvius *De Architectura* (2, 1) in reference to buildings (aedificiorum) and as a trope applied to speech with the meanings ‘structure’ and ‘skilful construction’ in Cicero’s *De Oratore* (3, 42, 167). Pokorny’s dictionary lists under entry no. 371 the root *dhabh-2* with the English meaning ‘proper’ and the German meaning ‘passend fügen, passend’ with reflexes in Armenian *darbin* (‘smith’), Latin *faber* (‘craftsman’, ‘artist’) together with a related adjective meaning ‘artful’, ‘skillful’ and the adverb *fabre* for ‘skillful’. The noun *fabrica* has the meanings ‘skillfulness’ and ‘workshop’. Reflexes in other languages include Gothic in *ga-daban* ‘occur’, ‘happen’ and Old Church Slavonic *dobrŭ* that is semantically equated with the ancient Greek terms ‘ἀγαθός’ and ‘καλός’. Entry 4112. *plastos* in *Strong’s Exhaustive Concordance of the Bible in Bible Hub* (retrieved 2025) has the lexical summary “*plastos*: fabricated, false, Original Word: πλαστός, Part of Speech: Adjective, Transliteration: *plastos*” with the meanings ‘feigned’ (KJV) and ‘false’ (NASB) and as the word’s etymological origin the verb *πλάσσω*. Its three meanings comprise ‘moulded’, ‘(by implication) artificial’, and ‘(figuratively) fictitious (false)’. The adjective related to entry 4112/*plastós* (‘plastic’, ‘fabricated’) occurs in the *Bible* (New Testament) only once in 2 *Pet* 2:3 for the description of deceptively molded words employed by false teachers and preachers in the expression ἐν πλεονεξίᾳ πλαστοῖς λόγοις ὑμῶς (GRK) translated as “they will exploit you with false words” (NAS). In a papyrus from the second century CE (POxy II. 237.viii.14, AD 186) the adjective *plastos* means ‘forged’ in the expression “not having immediately declared that the contract is forged”. The persons addressed in 2 *Peter* 2:1 correspond to the persons mentioned in 2 *Peter* 2:3: False prophets (ψευδοπροφῆται) and false teachers (ψευδοδιδάσκαλοι) who bring destruction (ἀπώλεια) (Westcott and Hort, 1881). Cicero in *De Oratore* (3, 42, 167) describes the fabrication (*fabricatio*) as a feature of the production of the speech (*oratio*) based on the rhetorical operations of transfer (*transductio*) and exchange (*immutatio*) illustrated by a sentence in which *Africans* is replaced by *Africa*. Cicero here states that what avoids obscurity (*obscuritas*) is a great embellishment (*magnum ornamentum*) of the speech (*oratio*),

since what comes into existence through this genre is called *riddles* ("etenim hoc fere genere fiunt ea, quae dicuntur aenigmata;"). Yet, this mode is not in the word (*in verbo modus hic*), but in the speech (*oratio*), that means, in the continuation of the words ("sed in oratione, id est, in continuatione verborum"). This transfer (*traductio*) and exchange (*immutatio*) do not constitute a certain fabrication (*quandam fabricationem*) in a word, [but in speech (*oratio*)] ("Ne illa quidem traductio atque immutatio in verbo quandam fabricationem habet [sed in oratione]"): Africa, a horrid land, tramples by terrible tumult ("Africa terribili tremit horrida terra tumultu;"). This example from the *Annals* written by Ennius Cicero demonstrates, by the replacement of *Africans* by *Africa*, the fabrication of speech.

This concept of fabrication can be transferred to AI applications like ChatGPT in terms of their automated operations as reflected in ChatGPT generated texts. When compared with existing scientific terminology in language studies and through a case study of an automated translation, the digitally constructed text is clearly a result of its technical environment. The fabrication occurs in AI generated processes through the construction of new output from given material, in analogy to the architectonic meaning of the word *fabrication*. Even ChatGPT describes its main technical feature, as previously shown in the prompt "Which computational methods and algorithms does ChatGPT use in order to analyze a digital text that was submitted to ChatGPT with a prompt to *analyze* its features as written speech?", as the *underlying architecture* of so-called *transformer neural networks*. Within this conceptual framework, the digitally manipulated text representing speech should be regarded as a construct formed in accordance with the transformative operations made possible by this imaginative underlying architecture due to machine learning models and natural language processing. The result of such fabrication, namely a false object or fabricated object, refers to the no longer existing reference between the digital object, which is generated through algorithmic probability and existing reality as a criterion of truth.

3.4. Modelling the Scenario of Speech and Speech-Like Digital Information: Borderlines between Human Speech and the Digital Representation of Human Speech

As for the process of communication of digital information in *ChatGPT*, originally human speech is transformed into digital data representing information without any label to indicate that it classifies as speech. Processed spoken or written speech in this digital environment not only constitutes a transformation into a different medium, but also represents a replication of particular features of human speech, basically the symbolic representation of an encoded meaning related to a text unit of spoken or written text (*Table 1*).

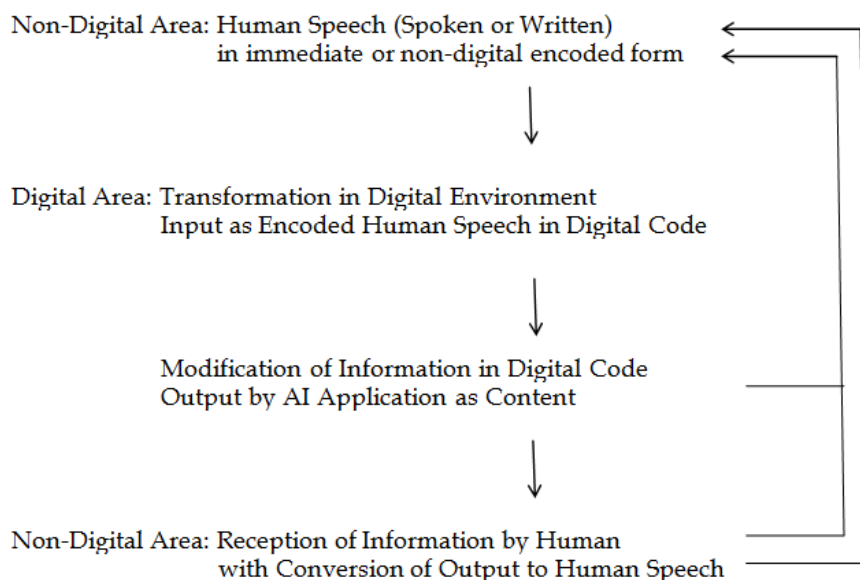


Table 1: *Model of Transformation for ChatGPT: Human Speech before Input and after Output of Digital Information in the Digital Environment*

We can locate the interface from the perspective of the user as the place where communication takes place. The condition of the "if" and the way in which digital output from the automation process of an AI application like ChatGPT enters the non-digital area as human speech depend on specific cases:. For example, a user can respond to the output commenting on it orally viewing it on the screen, or the user may process the output within the digital environment by writing a report. In both cases the digital content is conveyed into human speech. The processing of the model becomes circular in case speech-like digital data re-enter the communicative cycle through human reception and are subsequently reintroduced as digital data. In digital transformation processes produced data and information can be considered speech-like constructs representing speech (e.g., a digitally encoded monologue that is analyzed by the AI application with regard to its linguistic features). This production of representations of speech must be distinguished from a) the previous speech (monologue) before the digital transformation and b) the reception of digital data and information in human speech (e.g., a commentary of the results in spoken or written human speech).

Typologically, ChatGPT is a conversational AI, implicitly containing the communicative means of speech. As such, its technology entails Machine Learning (ML), including algorithms that learn continuously from datasets and user interactions, Natural Language Processing (NLP)/Natural Language Understanding (NLU), Natural Language Generation (NLG) producing human-like responses, and Automatic Speech Recognition (ASR) converting spoken words of a voice into text for AI processing. To describe the

practice of analyzing speech performed by ChatGPT we must not only consider the submitted speech as input undergoing a process of manipulation for an output, but also a complex conglomerate consisting of the digitalized speech, the prompt and the data of models in machine learning, and even the formal representation of conversational speech, all of which together constitute in the automated fabrication process a resulting text object characterized by a limited set of features of human speech.

Cited Works

ChatGPT: Prompts and Responses

ChatGPT Prompts: Translation of 2 Peter 2:3

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] Make a word list of the Greek words with an English translation of each word of the following text: καὶ ἐν πλεονεξίᾳ πλαστοῖς λόγοις ὑμᾶς ἐμπορεύσονται· οἷς τὸ κρίμα ἔκπαλαι οὐκ ἄργεῖ, καὶ ἡ ἀπώλεια αὐτῶν οὐ νυστάζει. December 17. <https://chatgpt.com>.

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] Make a translation of this text. December 17. <https://chatgpt.com>.

ChatGPT Prompts: Inquiry about Speech Processing by ChatGPT

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] How does *ChatGPT* define speech? December 17. <https://chatgpt.com>.

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] Which computational methods and algorithms does *ChatGPT* use in order to analyze a digital text that was submitted to *ChatGPT* with a prompt to *analyze* its features as written speech? December 17. <https://chatgpt.com>.

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] How does *ChatGPT* assess and classify a unit of spoken or written speech that is digitally presented to it? December 17. <https://chatgpt.com>.

ChatGPT; Haase, F.-A. [Prompt and response:] Which linguistic methods can *ChatGPT* use to analyze spoken or written speech that is digitally present? December 17. <https://chatgpt.com>.

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] Which linguistic methods can *ChatGPT* use to analyze a digital text? December 17. <https://chatgpt.com>.

ChatGPT; Haase, F.-A. (2025). [Prompt and response:] What can *ChatGPT* do to analyze a digital text? December 17. <https://chatgpt.com>.

References

*** (1879). *A Latin Dictionary. Founded on Andrews' Edition of Freund's Latin Dictionary, revised, enlarged, and in great part rewritten by Charlton T. Lewis and Charles Short.* Clarendon Press. <https://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0059%3Aentry%3Dfabricatio>.

- *** (1881). *New Testament*. Westcott & Hort (Eds). <https://biblehub.com>.
- *** (2012). *Collins English Dictionary. Complete & Unabridged*. Harper Collins Publishers. <https://www.dictionary.com>.
- *** (2025). *Oxford English Dictionary*. Oxford University Press. <https://doi.org/10.1093/OED/1077451145>.
- Cicero, M. T. (2025). De Oratore. In: W. L. Carey (Ed.). *The Latin Library*. <https://www.thelatinlibrary.com/cicero/oratore3.shtml>.
- Gordoni, M. (2025). The Slang Evolution: Why Gen Alpha Calls Each Other 'Chat'. *Movieguide*, June 30. <https://www.movieguide.org/news-articles/the-slang-evolution-why-gen-alpha-calls-each-other-chat.html#:~:text=%E2%80%9CChat%E2%80%9D%20refers%20to%20an%20audience,almost%20like%20its%20one%20person>.
- Insera, D. (2024). *Artificial Intelligence Regulation Threatens Free Expression*. Briefing Paper No. 178. July 16. <https://www.cato.org/briefing-paper/artificial-intelligence-regulation-threatens-free-expression#>.
- Pokorny, J. (2025). *Dictionary of the PIE Database*. <https://starlingdb.org/babel.php?lan=en>.
- Rafique, H., Nazeer, I., Rehman, J. (2024). Impact of ChatGPT on Language Evolution: A Linguistic Analysis. *Journal of Education and Social Studies*, 5(1), 56-68. https://www.researchgate.net/publication/378402344_The_Impact_of_ChatGPT_on_Language_Evolution_A_Linguistic_Analysis.
- Strong, J. (2025). 4112. plastos. In: *Strong's Exhaustive Concordance of the Bible*. Bible Hub. <https://biblehub.com/greek/4112.htm>.
- Tiwari, D., Arora, P. P., Geetey, Sh., Shrivastava, L., Shree, K. K. I., Melon, D. L., Nayagi, S. R. (2025). Language, Communication, and AI: The Role of Artificial Intelligence in Enhancing Public Speaking Skills. *International Journal of Environmental Sciences*, 11(23), 1401-1408. DOI:10.64252/vdrbfh86.
- Zhang, Z. (2025). AI-Powered Intelligent Speech Processing: Evolution, Applications and Future Directions. *International Journal of Advanced Computer Science and Applications*, 16(2), 918-928.