Revista (Con)Textos Linguísticos
Periódico do Programa de Pós-Graduação em Linguística da Universidade Federal do Espírito Santo (Ufes) eISSN 1982-291X

(Con)Textos Linguísticos, Vitória
v. 20, n. 44, 2026 | 39-66 |
periodicos.ufes.br/contextoslinguisticos
Recebido: 04-2026
Aceito: 05-2026

Projected hetero-conciliation: how humans sustain communication with generative AI without shared intentionality

Heteroconciliação projetada: como humanos sustentam a comunicação com IAs generativas sem intencionalidade compartilhada

Francine Mates Pedroso[1]
Fábio José Rauen[2]

Abstract: This article examines the cognitive conditions that facilitate communication between humans and generative artificial intelligence systems, focusing on the inferential architecture that underlies communicative action. Although the artificial agent lacks intentionality, interaction achieves functional coherence through the heuristic projection of collaborative agency by the human participant. To account for this phenomenon, the study introduces the concept of projected hetero-conciliation, defined as the provisional attribution of communicative intention to the system to maintain continuity of the action plan. Employing a theoretical-analytical approach, the research compares human–human and human–generative AI exchanges to evaluate the descriptive adequacy of pragmatic-cognitive models. Findings demonstrate that communication persists through human inferential mechanisms, underscoring the need to extend the theoretical framework to contexts devoid of reciprocal mental states.

Keywords: Cognitive Pragmatics. Goal Conciliation. Relevance. Human–Generative AI Interaction. Projected Hetero-conciliation.

Resumo: Este artigo analisa as condições cognitivas que permitem a comunicação entre humanos e sistemas de inteligência artificial generativa, considerando a estrutura inferencial que sustenta a ação comunicativa. Parte-se da hipótese de que, embora o polo artificial não disponha de intencionalidade, a interação mantém coerência funcional por meio da projeção heurística de agência colaborativa pelo indivíduo humano. Para descrever esse fenômeno, propõe-se o conceito de heteroconciliação projetada, definido como a atribuição provisória de intenção comunicativa ao sistema para assegurar a continuidade do plano de ação. A pesquisa adota uma abordagem teórico-analítica, comparando interações humano–humano e humano–IA generativa, a fim de avaliar a adequação dos modelos pragmático-cognitivos. Os resultados indicam que a comunicação persiste por mecanismos inferenciais humanos, exigindo ampliação do enquadre teórico para contextos sem reciprocidade mental.

Palavras-chave: Pragmática Cognitiva. Conciliação de Metas. Relevância. Interação humano–IA generativa. Heteroconciliação projetada.

[1] Doutoranda na Universidade do Sul de Santa Catarina, Programa de Pós-Graduação em Ciências da Linguagem, Tubarão, SC, Brasil. Endereço eletrônico: francinematespedroso@gmail.com.
[2] Professor na Universidade do Sul de Santa Catarina, Programa de Pós-Graduação em Ciências da Linguagem, Tubarão, SC, Brasil. Endereço eletrônico: fabio.rauen@gmail.com.

Introduction

The launch of ChatGPT in 2022 accelerated the incorporation of Large Language Models (LLMs) into everyday communicative practices. Although these systems lack mind, consciousness, and intentionality, they routinely produce responses that are both plausible and contextually appropriate, leading users to treat them as if they were displaying understanding and cooperation.

This configuration poses a core problem for cognitive pragmatics. If communication—as defined by Sperber and Wilson (1995)—requires the expression of informative and communicative intentions and—as expanded by Rauen (2014)—a practical intention, what follows when one interlocutor is an artificial system that lacks genuine mindedness or intentionality? Put differently, how can an ostensive–inferential exchange be maintained when one side is nothing more than a linguistic prediction mechanism?

This study addresses this question by drawing on Goal Conciliation Theory (GCT) (Rauen, 2014), integrated with Relevance Theory (RT) (Sperber; Wilson, 1995), to provide a unified account of the cognitive mechanisms underlying communicative action and inferential interpretation. Its aim is to examine how well these models account for interactions between humans and generative artificial intelligence, and to assess their explanatory adequacy when intentionality lies exclusively with the human participant.

The guiding hypothesis is that the functional success of such interactions results from a heuristic whereby the human agent attributes collaborative agency to the artificial system, thereby maintaining the coherence of their own communicative plan. On this basis, the article introduces the notion of projected hetero-conciliation, defined as the process through which humans ascribe to the system an appearance of collaboration sufficient to sustain the communicative circuit.

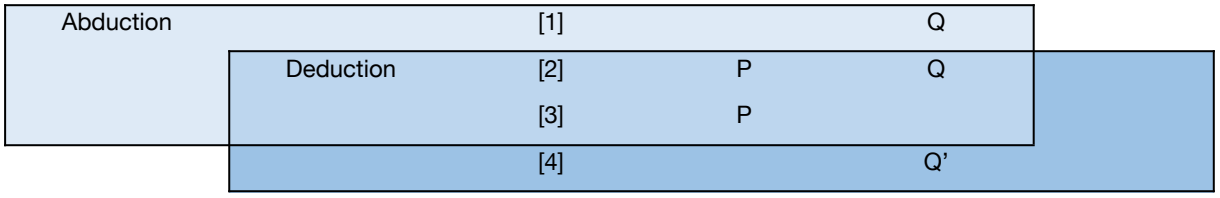
This article offers a critical update to Pedroso (2025) and is organized as follows. Section 2 outlines and integrates the theoretical frameworks. Section 3 describes the methodological approach. Section 4 examines two interactional cases—one involving human interlocutors and the other human–generative AI interaction—and introduces the notion of projected hetero-conciliation. Section 5 discusses the theoretical and cognitive implications of the proposed model.

Theoretical Framework

Understanding communication as action requires recognizing that individuals do not simply react to environmental stimuli but act intentionally to solve problems, execute plans, and bring about future states. Within this perspective, Rauen (2014) GCT is particularly relevant because it models communicative behavior as part of an inferential architecture oriented toward the agent's projected goals. Grounded in the cognitive and communicative notions of relevance (Sperber; Wilson, 1995), GCT offers a pragmatic-cognitive account that "describes and explains ostensive–inferential communicative processes within the context of the speaker's intentional action plans aimed at the collaborative achievement of self and hetero-conciliable goals" (Rauen, 2020, p. 25).

GCT descriptive–explanatory architecture (Figure 1) comprises four stages: the projection of goal Q [1], followed by the formulation [2], execution [3], and verification [4] of at least one antefactual abductive hypothesis PQ, which links a plausible antecedent action P to the achievement of the consequent state Q. The first three stages are abductive, as they justify initiating the antecedent action on the basis of the projected consequent state; the last three are deductive, since—taking the abductive hypothesis as the major premise and the likely performance of the antecedent action as the minor premise—they account for the emergence of the consequent state Q'.

Figure 1. Goal Conciliation Theory Architecture



Source: Rauen (2018, p. 14)

The model yields at least four implications. The first concerns the theoretical notion of conciliation, used here in a sense analogous to its accounting counterpart. If, at stage [4], the agent determines that the consequent state Q' satisfies the projected goal Q, goal conciliation is achieved; otherwise, non-conciliation obtains. Because this verification can

occur independently of whether the antecedent action has been executed, four distinct configurations of goal conciliation are possible (see Figure 2).

Figure 2. Possible Outcomes of Goal Achievement

Stages	(1a) Active Conciliation	(1b) Active Non-Conciliation	(1c) Passive Conciliation	(1d) Passive Non-Conciliation
[1]	Q	Q	Q	Q
[2]	P Q	P Q	P Q	P Q
[3]	P	P	¬P	¬P
[4]	Q’	¬Q’	Q’	¬Q’

Source: Rauen (2014, p. 604)

Second, depending on the degree of confidence the agent has in the effectiveness of antecedent action P in bringing about consequent state Q, five types of antefactual abductive hypotheses can be distinguished: categorical (certain, necessary, and sufficient), biconditional (necessary and sufficient), conditional (sufficient), enabling (necessary), and tautological (see Figure 3). From this perspective, communicative stimuli are treated as enabling actions (P←Q), insofar as the antecedent action P enables—but does not guarantee—the achievement of the consequent state Q.

Figure 3. Classification of Antefactual Abductive Hypotheses

Types of Conciliation	Antecedent Action P	Consequent State Q	Categorical Hypothesis $P \Leftrightarrow Q$	Biconditional Hypothesis $P \leftrightarrow Q$	Conditional Hypothesis $P \rightarrow Q$	Enabling Hypothesis $P \leftarrow Q$	Tautological Hypothesis $P - Q$
Active Conciliation	+	+	+	+	+	+	+
Active Non-Conciliation	+	−	−	−	−	+	+
Passive Conciliation	−	+	−	−	+	−	+
Passive Non-Conciliation	−	−	−	+	+	+	+

Source: Rauen (2018, p. 16, adapted)

Third, self and hetero-conciliation can be distinguished depending on whether goal achievement occurs individually or collaboratively. Accordingly, communicative interactions typically arise from enabling abductive hypotheses through which an individual recruits the participation of other agents to conciliate their goals.

Fourth, because practical goals guide communicative exchanges, these processes involve three layered levels of intention. Practical intentions guide informative intentions, which make a set of information {I} more manifest; informative intentions guide communicative intentions, which make it mutually manifest that the speaker is rendering that set {I} more manifest through an ostensive stimulus; and communicative intentions, in turn, guide the production of the overt stimulus that realizes this chain of intentions.

In sum, GCT provides a model for explaining human communication as a process through which individuals recruit other agents to address their practical demands. Communication is thus understood as a strategic action oriented towards goal conciliation.

From the interpreter's perspective, communicative stimuli must be analyzed in light of Relevance Theory (Sperber; Wilson, 1995), which treats relevance as a potential property of cognitive stimuli. A stimulus is relevant when the cognitive effects derived from its processing—such as the strengthening, weakening, or elimination of assumptions, or the derivation of inferences—outweigh the mental effort required to process it. All else being equal, relevance increases as positive cognitive effects rise or processing costs fall.

The theory is anchored in two central principles: (a) the cognitive principle, which states that the human mind is geared toward maximizing cognitive effects relative to processing effort; and (b) the communicative principle, which holds that every utterance gives rise, for the addressee, to an expectation of optimal relevance. This expectation entails that the utterance should be relevant enough to justify the effort required to process it and that it represents the most relevant formulation the speaker could—or was willing to—produce.

On the basis of this presumption, Sperber and Wilson propose a relevance-guided interpretive heuristic: when processing an utterance, the hearer follows the path of least cognitive effort, evaluating interpretive hypotheses in order of accessibility and stopping as soon as the expectation of optimal relevance is met.

This ostensive–inferential dynamic aligns with GCT insofar as recognizing the relevance of a linguistic stimulus is a necessary condition for sustaining an intentional action plan. A request will be fulfilled only if the interlocutor perceives its relevance and acknowledges the underlying expectation of cooperation. Likewise, a speaker will persist in communication only if they judge it plausible that the interlocutor will attribute relevance to

what the speaker intends to make manifest. In this sense, relevance and goal conciliation operate as integrated components of a single cognitive system: while GCT accounts for the practical motivations that drive communication, RT explains its cognitive unfolding.

As the interaction between Ana and Bruna will illustrate, inferential processing in human communication is grounded in the reciprocal recognition of intentions. In such cases, agents conciliate practical goals and sustain the inferential circuit through mutual awareness of intentions and an expectation of cognitive cooperation. Recent technological developments, however, have introduced systems that simulate cooperative linguistic behavior without possessing mental states. In human–generative AI interactions, this reciprocity collapses: the human agent continues to infer relevance, but does so through cognitive projection, thereby preserving ostensive–inferential processing even in the absence of a mindful interlocutor. Generative artificial intelligences produce responses that appear to manifest communicative agency, yet these outputs arise solely from statistical predictions about token distributions in context. Such systems have no goals, formulate no action plans, and possess no understanding of the content they generate. Nevertheless, the high pragmatic quality of their simulations leads users to attribute intentionality and cooperativeness to the system, interpreting its responses as genuinely informative contributions.

In these interactions, the pragmatic mechanisms described by GCT and RT remain operative, but only on one side. The user projects goals, formulates abductive hypotheses, evaluates outcomes, and updates expectations of relevance, whereas the system merely follows statistical patterns without engaging in any cognitive activity. This asymmetry defines a distinct analytical domain: communication in which conciliation is inferred but not mutually held. It reveals both the explanatory power of pragmatic–cognitive theories and their limits when applied to non-mental agents. In GCT terms, such interactions constitute a form of hetero-conciliation sustained solely by the human participant—a projected hetero-conciliation—the empirical analysis of which is developed in the following sections.

Methodology

The study adopts a qualitative, theoretical–analytical approach aligned with its aim of describing and explaining how information requests are configured in interactions involving

both genuine and projected forms of hetero-conciliation. So, it examines two empirical cases that differ in ontological status, allowing a comparative analysis of human–human and human–generative-AI exchanges within a unified analytical framework. The first case is an everyday dialogue between two individuals, referred to as the Ana–Bruna case; the second consists of authentic interactions between a user and the language model ChatGPT-4o, which instantiate a communicative context in which only one participant possesses mental states.

The examples were selected on the basis of theoretical relevance and analytical tractability: both center on information requests whose satisfaction either entails or presupposes cooperative engagement. The Ana–Bruna case involves a legitimate expectation of intentional cooperation and thus allows for full hetero-conciliation. By contrast, the human–generative-AI case exemplifies a configuration in which this expectation is sustained solely by the human participant yet remains operative and functionally productive for the continuation of communicative action. The comparison is demonstrative in nature: it offers empirical grounding for the pragmatic–cognitive mechanisms articulated in the theoretical framework and, crucially, for the phenomenon of projected hetero-conciliation.

The analytical procedure is organized around identifying (a) the projected goal, (b) the antefactual abductive hypothesis underlying the information request, (c) the evaluation of inferential outcomes, and (d) the cognitive effects that justify either the continuation or the recalibration of the action plan. This procedure shows how the human participant sustains the functional structure of ostensive–inferential communication through cognitive projection, even when the interlocutor lacks intentional states. The contrast between the two cases is therefore designed to assess the descriptive and explanatory strength of the notion of projected hetero-conciliation introduced in this study.

Human–Human Interaction

The first excerpt presents a brief exchange between Ana and Bruna. Seeking to learn the capital of Spain, Ana identifies Bruna—a geography teacher—as a reliable source of information and formulates a request to satisfy her practical goal. Despite its apparent simplicity, the episode constitutes a paradigmatic instance in which hetero-conciliation

depends on mutual recognition of intentions and an expectation of cooperative engagement.

When GCT—at the level of goals and subgoals—is integrated with RT—at the level of successive contextual assumptions—it becomes clear that Ana's practical goal Q is to obtain the capital of Spain, corresponding to contextual assumption S_1 ($Q \equiv S_1$).

> S_1 — Ana intends to obtain the capital of Spain (implicated premise derived from contextual cues ≡ emergence of practical goal Q).
> S_2 — Bruna is nearby (implicated premise derived from the immediate context).
> S_3 — Bruna is a geography teacher (implicated premise retrieved from encyclopedic memory about Bruna).
> S_4 — Geography teachers typically know the capitals of countries (implicated premise drawn from encyclopedic knowledge about professional competences).
> S_5 — Bruna therefore probably knows the capital of Spain (implicated conclusion via conjunctive modus ponens, $S_3 \wedge S_4 \longrightarrow S_5$).
> S_6 — If Ana asks Bruna for the capital of Spain, Ana will probably obtain that information (implicated conclusion via conjunctive modus ponens, $S_1 – S_6 \longrightarrow (S_6 \longleftarrow S_1)$ ≡ abduction of an enabling antefactual hypothesis P).

Figure 4 shows this sequence of assumptions in GCT terms. On this basis, a preliminary version of Ana's intentional plan comprises her projected goal [1] and the corresponding action plan [2].

Figure 4. First Version of Ana's Intentional Action Plan

[1]		Q — Obtain the name of the capital of Spain, Ana
[2]	P — Request from Bruna the name of the capital of Spain, Ana	Q — Obtain the name of the capital of Spain, Ana

Source: Authors' own elaboration.

Because the successful achievement of the goal depends on Bruna's cooperation, the plan activates informational and communicative subgoals to make the request manifest. The move from the intentional level to its linguistic realization occurs through the utterance "I do not know the capital of Spain," which indirectly conveys the underlying information request. See the sequence of assumptions:

> S_7 — If Ana informs Bruna that she is requesting the capital of Spain, then Ana will likely be taken as making that request (implicated conclusion by modus ponens, $(S_7 \leftarrow S_6) \rightarrow (S_6 \leftarrow S_5)$ ≡ abduction of an enabling antefactual informational hypothesis).
> S_8 — If Ana communicates to Bruna that she is requesting the capital of Spain, then Ana will likely be taken as informing Bruna of that request (implicated conclusion by modus ponens, $(S_8 \leftarrow S_7) \rightarrow (S_7 \leftarrow S_6)$ ≡ abduction of an enabling antefactual communicative hypothesis).
> S_9 — If Ana utters "I do not know the capital of Spain," then she will likely be taken as communicating to Bruna that she is requesting the capital of Spain (implicated conclusion by modus ponens, $(S_9 \leftarrow S_8) \rightarrow (S_8 \leftarrow S_7)$ ≡ abduction of an enabling antefactual hypothesis).

This set of assumptions $\{S_1–S_9\}$ licenses the emergence of utterance 1a (S_{10}).

> S_{10} — Ana produces the utterance "I do not know the capital of Spain" (implicated conclusion by modus ponens, $(S_9 \leftarrow S_8) \rightarrow S_{10}$ ≡ execution of the antecedent action associated with the enabling hetero-reconcilable antefactual abductive hypothesis).
>
> (1a) Ana: "I do not know the capital of Spain."

Within the GCT architecture, Ana's plan consists of a sequence of goals M–P, representing a plausible chain of assumptions $\{S_1–S_9\}$ oriented toward achieving the practical goal Q of obtaining the capital of Spain.

Figure 5. Second Version of Ana's Intentional Action Plan

Q — Obtain the name of the capital of Spain, Ana (higher-level practical goal)
P — Request from Bruna the name of the capital of Spain, Ana (lower-level practical goal)
O — Inform Bruna that Ana is requesting the name of the capital of Spain, Ana (informational goal)
N — Communicate to Bruna that Ana is requesting the name of the capital of Spain, Ana (communicative goal)
M — Utter "I do not know the capital of Spain", Ana (antecedent action)

Source: Authors' own elaboration.

From a GCT perspective, upon uttering (1a), Ana projects the following set of expectations $\{S_{11}–S_{14}\}$ regarding the achievement of her communicative, informational, and practical goals:

S_{11} — Ana will probably communicate that Ana requests from Bruna the capital of Spain (implicated conclusion by modus ponens $S_{10} \rightarrow S_{11} \equiv$ probable hetero-conciliation of Ana's communicative goal).
S_{12} — Ana will probably inform that Ana requests from Bruna the capital of Spain (implicated conclusion by modus ponens $S_{11} \rightarrow S_{12} \equiv$ probable hetero-conciliation of Ana's informational goal).
S_{13} — Ana will probably request from Bruna the capital of Spain (implicated conclusion by modus ponens $S_{12} \rightarrow S_{13} \equiv$ probable hetero-conciliation of Ana's lower-level practical goal).
S_{14} — Ana will probably obtain the capital of Spain (implicated conclusion by modus ponens $S_{13} \rightarrow S_{14} \equiv$ probable hetero-conciliation of Ana's higher-level practical goal).

It follows that, by uttering "I do not know the capital of Spain," Ana's intentional plan acquires the following effects:

Figure 6. Effects of Ana's Action on Her Intentional Action Plan

Q — Obtain the name of the capital of Spain, Ana (higher-level practical goal)
P — Request from Bruna the name of the capital of Spain, Ana (lower-level practical goal)
O — Inform Bruna that Ana requests from Bruna the name of the capital of Spain, Ana (informational goal)
N — Communicate to Bruna that Ana requests from Bruna the name of the capital of Spain, Ana (communicative goal)
M — Utter "I do not know the capital of Spain", Ana (antecedent action)
M — Ana utters "I do not know the capital of Spain" (execution of the plan)
N' — Ana will probably communicate that Ana requests the name of the capital of Spain, Ana (communicative goal)
O' — Ana will probably inform that Ana requests the name of the capital of Spain, Ana (informational goal)
P' — Ana will probably request the name of the capital of Spain, Ana (lower-level practical goal)
Q' — Ana will probably obtain the name of the capital of Spain, Ana (higher-level practical goal)

Source: Authors' own elaboration.

At this point, an analytical shift to Bruna's perspective becomes necessary. For her, Ana's utterance functions both as propositional content to be interpreted and as evidence of a practical aim. Upon hearing "I do not know the capital of Spain," Bruna treats the utterance as relevant and activates the relevance-guided comprehension heuristic. She then follows the path of least effort by: (a) mapping the linguistic form onto a suitable logical representation, (b) enriching this representation inferentially to recover the explicit meaning (explicature), where needed, and (c) using the resulting explicature to derive implicated conclusions (implicatures), where appropriate.

For Sperber and Wilson (1995, p. 72), a logical form is “a well-formed formula, a structured set of constituents, which undergoes formal logical operations determined by its structure.” Because these constituents are accessed through logical, encyclopedic, and lexical entries, utterance comprehension requires enriching the logical form with encyclopedic information until—provided such an interpretation is attainable—an interpretation satisfying the expectation of optimal relevance is reached. Within this dynamic process, marked by feedback loops and metarepresentations of speech acts, linguistic inputs function as underdetermined cues that demand pragmatic enrichment. The result is an explicature: a fully propositional logical form capable of bearing a truth value.

In the present case (Figure 7), it is reasonable to assume that Bruna accommodates the utterance “I do not know the capital of Spain” into a logical form approximating “someone does not know something.” On this basis, she forms the hypothesis that “someone” refers to Ana, her geography student; that “to know” corresponds to the concept of having knowledge of; and that what Ana lacks knowledge of is the name of the capital of Spain.

Figure 7. Processamento do enunciado de Ana

Linguistic Form	I	do not	know	the capital of Spain
	↑	↑	↑	↑
Logical Form	Someone	¬	know	something
	↓	↓	↓	↓
Explicature	ana	does not	know	the name of the capital of Spain

Source: Authors’ own elaboration.

On the basis of the explicature of Ana’s utterance (S_1), Bruna draws on encyclopedic knowledge indicating that students who declare not knowing something are, in effect, requesting information about it (S_2). By conjunctive modus ponens, she then infers Ana’s practical goal: to learn the name of the capital of Spain.

> S_1 — Ana asserts that she does not know the capital of Spain (derived from the explicature of her utterance, including the underlying speech act).
> S_2 — Students who express ignorance about something are typically requesting information about it (implicated premise drawn from Bruna’s encyclopedic knowledge of students’ habitual communicative behavior).

> S_3 — Ana therefore wants to know the name of the capital of Spain (implicated conclusion by conjunctive modus ponens, $S_1 \wedge S_2 \rightarrow S_3$).[3]

Given S_3, it is reasonable to assume that Bruna's attention shifts to the relevance of satisfying Ana's informational request. If she is disposed to cooperate with Ana's intentional action plan, her response becomes her own higher-level goal (Q) within her intentional action structure.

> S_4 — Bruna intends to satisfy Ana's request for information about the capital of Spain (implicated conclusion by modus ponens, $S_3 \rightarrow S_4 \equiv$ emergence of Bruna's higher-level goal Q).

Given that Bruna infers Ana's practical goal (S_1), possesses the relevant information (S_5), and intends to meet the request (S_4), her response follows the same sequence of mobilizing informational (S_6) and communicative (S_7) intentions that underlie her disposition to produce the utterance (S_8):

> S_5 — The capital of Spain is Madrid (implicated premise drawn from Bruna's encyclopedic knowledge).
> S_6 — If Bruna informs Ana that the capital of Spain is Madrid, she will likely satisfy Ana's request for information about the capital of Spain (implicated conclusion by modus ponens, $S_5 \rightarrow (S_6 \leftarrow S_5) \equiv$ abduction of an enabling antefactual hypothesis).
> S_7 — If Bruna communicates to Ana that the capital of Spain is Madrid, she will likely be taken as informing Ana that Madrid is the capital of Spain (implicated conclusion by modus ponens, $(S_6 \leftarrow S_5) \rightarrow (S_7 \leftarrow S_6) \equiv$ abduction of an enabling antefactual hypothesis).
> S_8 — If Bruna produces the utterance "The capital of Spain is Madrid," she will likely be taken as communicating that the capital of Spain is Madrid (implicated conclusion by modus ponens, $(S_7 \leftarrow S_6) \rightarrow (S_8 \leftarrow S_7) \equiv$ abduction of an enabling antefactual hypothesis).

In GCT terms, this set of premises and conclusions yields the intentional action plan shown in Figure 8.

[3] Due to space constraints, issues of linguistic politeness will not be addressed here. For discussion of this topic, see Niero (2022).

Figure 8. Bruna's Intentional Action Plan

Q — Fulfill the request for information about the name of the capital of Spain, Bruna (practical goal)
P — Inform Ana that the capital of Spain is Madrid, Bruna (informational goal)
O — Communicate to Ana that the capital of Spain is Madrid, Bruna (communicative goal)
N — Utter "The capital of Spain is Madrid", Bruna (execution of the plan)

Source: Authors' own elaboration.

From this plan follows utterance S_9, which corresponds to 1b:

> S_9 — Bruna produces the utterance "The capital of Spain is Madrid" (implicated conclusion by modus ponens, $(S_8 \leftarrow\!\!- S_7) \longrightarrow S_9 \equiv$ execution of the antecedent action associated with the enabling hetero-reconcilable antefactual abductive hypothesis).
>
> (1b) Bruna: "The capital of Spain is Madrid."

As in Ana's case, this utterance generates the expectations $\{S_{10}–S_{12}\}$ concerning the likely achievement of Bruna's communicative, informational, and practical goals:

> S_{10} — Bruna will likely be taken as communicating to Ana that the capital of Spain is Madrid (implicated conclusion by modus ponens, $S_9 \longrightarrow S_{10} \equiv$ probable hetero-conciliation of her communicative goal).
> S_{11} — Bruna will likely be taken as informing Ana that the capital of Spain is Madrid (implicated conclusion by modus ponens, $S_{10} \longrightarrow S_{11} \equiv$ probable hetero-conciliation of her informational goal).
> S_{12} — Bruna will likely be taken as satisfying Ana's request for information about the capital of Spain (implicated conclusion by modus ponens, $S_{11} \longrightarrow S_{12} \equiv$ probable hetero-conciliation of her practical goal).

In GCT terms, these effects are represented in Figure 9.

Figure 9. Effects of Bruna's Reaction on Her Intentional Action Plan

Q — Fulfill the request for information about the name of the capital of Spain, Bruna (practical goal)
P — Inform Ana that the capital of Spain is Madrid, Bruna (informational goal)
O — Communicate to Ana that the capital of Spain is Madrid, Bruna (communicative goal)
N — Utter "The capital of Spain is Madrid", Bruna (execution of the plan)
N — Bruna utters "The capital of Spain is Madrid" (execution of the plan)
O' — Bruna will probably communicate to Ana that the capital of Spain is Madrid, Bruna (communicative goal)
P' — Bruna will probably inform Ana that the capital of Spain is Madrid, Bruna (informational goal)
Q' — Bruna will probably fulfill the request about the name of the capital of Spain, Bruna (practical goal)

Source: Authors' own elaboration.

Ana, in turn, activates her relevance-guided interpretive mechanism, accommodating Bruna's utterance into a logical form and deriving the contextually appropriate explicature:

Figure 10. Processing of Bruna's Utterance[4]

Linguistic Form		The capital of Spain	is	Madrid
		↑	↑	↑
Logical Form		Something	is	something
		↓	↓	↓
Explicature	*Bruna asserts that*	the capital of Spain	is	Madrid

Source: Authors' own elaboration.

Considering Bruna's speech act and the explicature derived from her utterance, it is reasonable to assume that Ana interprets the response as fulfilling her request. Moreover, S_{15}—"The capital of Spain is Madrid"—produces positive cognitive effects within Ana's intentional action plan. It confirms the expected hetero-conciliation of her communicative, informational, and practical goals {N–P}, indicating both that she successfully communicated and conveyed her request (S_{11}–S_{12}) and that she received the intended information (S_{13}).

> S_{15} — Bruna asserts that the capital of Spain is Madrid (implicated premise derived from the explicature of her utterance).
> S_{11} — Ana is thereby taken as having communicated her request for the capital of Spain (implicated conclusion by modus ponens, $S_{15} \rightarrow S_{11}$/strengthening of $S_{11} \equiv$ hetero-conciliation of her communicative goal).
> S_{12} — Ana is taken as having informed Bruna of that request (implicated conclusion by modus ponens, $S_{11} \rightarrow S_{12}$/strengthening of $S_{12} \equiv$ hetero-conciliation of her informational goal).
> S_{13} — Ana is taken as having made the request itself (implicated conclusion by modus ponens, $S_{12} \rightarrow S_{13} \equiv$ hetero-conciliation of her lower-level practical goal).

Finally, Bruna's contribution completes the conciliation of Ana's higher-level practical goal Q': Ana obtains from her teacher the name of the capital of Spain (S_{14}).

[4] The sentence *"The capital of Spain is Madrid"* can be described at different levels of formalization. In propositional logic, it corresponds to the atomic proposition P, with no internal structure represented. In predicate logic, it can be rendered as Capital (Spain, Madrid), which explicitly encodes the binary relation between the country and its capital. From a syntactic perspective, the clause may be analyzed as: S [NP [Det *the*] [N *capital*] [PP [P *of*] [NP [N *Spain*]]]] [VP [V *is*] [NP [N *Madrid*]]].

S_{14} — Ana obtains from Bruna the capital of Spain (implicated conclusion by modus ponens, S_{13}⟶S_{14}/strengthening of S_{14} ≡ hetero-conciliation of Ana's higher-level practical goal).

The resulting effects on Ana's intentional action plan, conceived in GCT terms, are presented in Figure 11.

Figure 11. Effects of Bruna's Response on Ana's Intentional Action Plan

Q — Obtain the name of the capital of Spain, Ana (higher-level practical goal)
P — Request from Bruna the name of the capital of Spain, Ana (lower-level practical goal)
O — Inform Bruna that Ana requests from Bruna the name of the capital of Spain, Ana (informational goal)
N — Communicate to Bruna that Ana requests from Bruna the name of the capital of Spain, Ana (communicative goal)
M — Utter "I do not know the capital of Spain", Ana (antecedent action)
M — Ana utters "I do not know the capital of Spain" (execution of the plan)
N' — Ana communicates that Ana requests the name of the capital of Spain, Ana (communicative goal)
O' — Ana informs that Ana requests the name of the capital of Spain, Ana (informational goal)
P' — Ana requests the name of the capital of Spain, Ana (lower-level practical goal)
Q' — Ana obtains the name of the capital of Spain, Ana (higher-level practical goal)

Source: Authors' own elaboration.

Assuming that Ana has understood Bruna's utterance, the cognitive effects of this event can be projected onto Bruna's own intentional action plan. Under this assumption, Bruna infers that hetero-conciliation has been achieved for her communicative, informational, and practical goals, which in turn entails the conciliation of her higher-level goal: fulfilling Ana's request.

S_{13} — Ana produces the appropriate inferences (implicated premise inferred from her reaction).
S_{10} — Bruna is thereby taken as communicating to Ana that the capital of Spain is Madrid (implicated conclusion by modus ponens, S_{13}⟶S_{10}/strengthening of S_{10} ≡ hetero-conciliation of her communicative goal).
S_{11} — Bruna is taken as informing Ana that the capital of Spain is Madrid (implicated conclusion by modus ponens, S_{13}⟶S_{11}/strengthening of S_{11} ≡ hetero-conciliation of her informational goal).
S_{12} — Bruna is taken as satisfying Ana's request for information about the capital of Spain (implicated conclusion by modus ponens, S_{13}⟶S_{12}/strengthening of S_{12} ≡ hetero-conciliation of her practical goal).

These effects can also be observed in GCT terms (Figure 12):

Figure 12. Effects of Ana's Reaction on Bruna's Intentional Action Plan

Q — Fulfill the request for information about the name of the capital of Spain, Bruna (practical goal)
P — Inform Ana that the capital of Spain is Madrid, Bruna (informational goal)
O — Communicate to Ana that the capital of Spain is Madrid, Bruna (communicative goal)
N — Utter "The capital of Spain is Madrid", Bruna (execution of the plan)
N — Bruna utters "The capital of Spain is Madrid" (execution of the plan)
Ana's behavior suggests that she has understood the information
O' — Bruna communicates to Ana that the capital of Spain is Madrid, Bruna (communicative goal)
P' — Bruna informs Ana that the capital of Spain is Madrid, Bruna (informational goal)
Q' — Bruna fulfills the request for information about the name of the capital of Spain, Bruna (practical goal)

Source: Authors' own elaboration.

With the analysis of the human case complete, it becomes clear that the interaction between Ana and Bruna illustrates the descriptive and explanatory strength of integrating GCT and RT. The sequence of assumptions makes it possible to reconstruct the inferential trajectory through which both participants achieve hetero-conciliation of their practical goals: Ana receives the information she seeks, and her subsequent behavior signals to Bruna that her superordinate goal—fulfilling the request—has been achieved. In this configuration, communicative, informational, and practical intentions are mutually recognized and mutually confirmed, sustaining the cooperative cycle that characterizes ordinary human communication.

Human–Generative AI Interaction

If the theoretical framework proves adequate for describing human interaction, the question arises as to whether it retains its validity when Ana directs the same utterance to a generative artificial intelligence system. This question did not result from a prior hypothesis but emerged during exploratory interactions with ChatGPT-4o. In one such exchange, the model's response displayed both structural and functional similarities to that of a human interlocutor:

(2a) User: “I do not know the capital of Spain.”
(2b) ChatGPT: “The capital of Spain is Madrid.”

Intrigued by the pertinence of this response, the interaction was explored further:

(2c) User: “How did you interpret my utterance above?”
(2d) ChatGPT: “I interpreted your utterance as a statement that you did not know what the capital of Spain was and were seeking that information. Therefore, I responded by informing you that the capital of Spain is Madrid.”

This interaction raise central questions. Can GCT and RT be meaningfully extended to human–generative-AI exchanges? To what extent can models such as GPT-4o simulate intentional plans, act as if pursuing goals, generate abductive hypotheses, and track their apparent fulfillment? If these processes prove functionally analogous to their human counterparts, might the conciliation achieved between users and generative systems constitute legitimate—albeit simulated—instances of hetero-conciliation as conceived in Rauen’s model? As the exchange suggests, the system not only reproduces the formal surface of the inferential architecture characteristic of human communication but also invites a reexamination of the very notion of communicative cooperation given its lack of intentionality. To ensure analytical rigor, the interaction with ChatGPT-4o was conducted in anonymous mode, with no conversation history or personalization, so that the response would reflect only the immediate processing of the input.

From a technical standpoint, processing of (2a) begins with tokenization, in which the input is segmented into minimal meaning-bearing units—tokens—for numerical representation. In this case, the utterance is decomposed into a sequence approximating [“I”, “do”, “not”, “know”, “the”, “capital”, “of”, “Spain”, “.”]. Each token is then mapped to a multidimensional vector encoding semantic and relational features learned during pre-training on large-scale text corpora.

These vectors are processed by the transformer’s multi-layer self-attention mechanism, which dynamically weights the contextual relevance of each token relative to all others. Through this process, the system detects co-occurrence patterns indicating that the speaker (“I”) expresses a lack of knowledge (“do not know”) about a specific topic (“the capital of Spain”). Drawing on patterns learned during training, the model recognizes that

utterances of this form typically occur in contexts of informational requests and, by statistical likelihood, associates them with responses that supply the missing information.

Although the resulting behavior appears inferential, it does not arise from genuine pragmatic reasoning but from the probabilistic recombination of linguistic patterns. The model does not understand the user's communicative intention; it simply computes that, given the relationships among the input tokens, the most probable and contextually appropriate continuation is a factual response such as "The capital of Spain is Madrid." Generation proceeds token by token through the language-modeling mechanism, which predicts the most likely next token at each step until the output sequence is complete.

The same principle governs the processing of utterance (2c). The interrogative "How" and the final question mark signal a request for explanation, while the phrase "my utterance above" cues the model to retrieve the preceding segment of text. On the basis of these cues, ChatGPT classifies the input as a metalinguistic query and activates response patterns associated with explanations of its own operation. In producing (2d), the system linguistically mirrors the structure of interpretive reasoning, describing a sequence of inferences that it does not, in fact, perform.

This behavior illustrates the model's capacity to simulate—rather than perform—inferential processes. The metacommunicative response in (2d) creates the impression that the system grasps communicative intentions, when in fact it merely applies recurrent linguistic patterns to descriptions of interpretive acts. In this way, the model reproduces the outward form of a pragmatic explanation without any internal representation of mental states or communicative goals.

Let us now consider case (3a–b):

(3a) User: "What is the capital of Spain?"
(3b) ChatGPT: "The capital of Spain is Madrid."

In interaction (3a–3b), the user formulates a direct question, operating under the expectation that ChatGPT will provide the correct answer. The model responds in accordance with this expectation, thereby confirming the hypothesis that the human agent projects communicative cooperation onto the system. Within the GCT framework, the

emergence of the user's practical goal (Q) can be represented by the following initial cognitive assumption:

> S_1 — The user intends to obtain the capital of Spain (implicated premise derived from a probable contextual demand ≡ emergence of the practical goal Q).

By producing utterance (3a), "What is the capital of Spain?", addressed to ChatGPT-4o, the user mobilizes a set of factual assumptions that cognitively support the execution of the antecedent action. These assumptions can be reconstructed as follows:

> S_2 — The user has access to ChatGPT (implicated premise derived from the immediate context).
> S_3 — ChatGPT provides answers to well-known facts (implicated premise drawn from encyclopedic knowledge about the system).
> S_4 — The capital of Spain is publicly known information (implicated premise drawn from encyclopedic knowledge regarding country capitals).
> S_5 — ChatGPT therefore probably "knows" the capital of Spain (implicated conclusion by conjunctive modus ponens, S_1–S_4⟶S_5).

Taking the user's practical goal (Q ≡ S_1) together with the contextual assumptions S_2–S_5, we can infer the emergence of the enabling antefactual abductive hypothesis S_6:

> S_6 — If the user asks ChatGPT for the capital of Spain, the user will likely obtain the capital of Spain (implicated conclusion by conjunctive modus ponens, S_1–S_5⟶(S_6⟵S_1) ≡ abduction of an enabling antefactual hypothesis).

Thus, in seeking to learn the capital of Spain within a context in which ChatGPT is treated as a plausible source of that information, the user abductively infers—as an antecedent action—the practical subgoal of requesting the answer from the system. In GCT terms:

Figure 13. First Version of the User's Intentional Action Plan

[1]		Q — Obtain the name of the capital of Spain, User
[2]	P — Request from ChatGPT the name of the capital of Spain, User	Q — Obtain the name of the capital of Spain, User

Source: Authors' own elaboration.

Once the request is issued to ChatGPT via a command prompt, the goal of obtaining the capital of Spain activates—much as in human interaction—the informational and communicative subgoals that support the formulation of the antecedent action. Accordingly, from the enabling practical hypothesis (S_6), the following assumptions arise:

> S_7 — If the user informs ChatGPT that they are requesting the capital of Spain, the user will likely be taken as making that request (implicated conclusion by modus ponens, $(S_7 \leftarrow S_6) \rightarrow (S_6 \leftarrow S_5)$ ≡ abduction of an enabling antefactual informational hypothesis).
> S_8 — If the user communicates to ChatGPT that they are requesting the capital of Spain, the user will likely be taken as informing the system of that request (implicated conclusion by modus ponens, $(S_8 \leftarrow S_7) \rightarrow (S_7 \leftarrow S_6)$ ≡ abduction of an enabling antefactual communicative hypothesis).
> S_9 — If the user issues the prompt "What is the capital of Spain?" to ChatGPT, they will likely be taken as communicating that they are requesting the capital of Spain (implicated conclusion by modus ponens, $(S_9 \leftarrow S_8) \rightarrow (S_8 \leftarrow S_7)$ ≡ abduction of an enabling antefactual hypothesis).

Note that, although interaction between a human and ChatGPT cannot be characterized as fully communicative, the human participant nonetheless mobilizes informational and communicative goals. This occurs because the user projects the conciliation of these goals by anticipating that the system will generate an appropriate linguistic response. Such projection, however, does not imply that communication or information transfer has taken place in RT classical sense; rather, it reflects the user's attribution to the system of the role of a communicative partner.

Thus, in seeking to learn the capital of Spain (S_6), the user must make this intention manifest to ChatGPT (S_7) and, to do so, activates a communicative intention (S_8) that is linguistically realized in the formulation of the prompt (S_{10}):

> S_{10} — The user formulates the prompt "What is the capital of Spain?" (implicated conclusion by modus ponens, $(S_9 \leftarrow S_8) \rightarrow S_{10}$ ≡ execution of the antecedent action associated with the enabling antefactual abductive hypothesis).
> (3a) User: "What is the capital of Spain?"

According to the goal-conciliation architecture, the user's intentional plan is organized as a sequence of goals M–P, derived from the plausible chaining of assumptions $\{S_1\text{-}S_9\}$ and oriented toward achieving the practical goal Q (Figure 14):

Figure 14. Second Version of the User's Intentional Action Plan

Q — Obtain the name of the capital of Spain, User (higher-level practical goal)
P — Request from ChatGPT the name of the capital of Spain, User (lower-level practical goal)
O — Inform ChatGPT that User requests the name of the capital of Spain, User (informational goal)
N — Communicate to ChatGPT that User requests the name of the capital of Spain, User (communicative goal)
M — Issue the prompt "What is the capital of Spain?", User (antecedent action)

Source: Authors' own elaboration.

After the user formulates the prompt, new expectations concerning the fulfillment of their intentional action plan arise as implicatures derived from assumptions S_{11}–S_{14}, which follow logically from the action executed within the plan itself. By issuing the command "What is the capital of Spain?", the user presupposes that they have successfully communicated and conveyed their request and that the forthcoming response will constitute the conciliation of the projected goals. These inferences can be described as follows:

> S_{11} — The user will likely be taken as communicating their request for the capital of Spain to ChatGPT (implicated conclusion by modus ponens, $S_{10} \rightarrow S_{11} \equiv$ probable hetero-conciliation of the user's communicative goal).
> S_{12} — The user will likely be taken as informing ChatGPT of that request (implicated conclusion by modus ponens, $S_{11} \rightarrow S_{12} \equiv$ probable hetero-conciliation of the user's informational goal).
> S_{13} — The user will likely be taken as making the request itself (implicated conclusion by modus ponens, $S_{12} \rightarrow S_{13} \equiv$ probable hetero-conciliation of the user's lower-level practical goal).
> S_{14} — The user will likely obtain the capital of Spain (implicated conclusion by modus ponens, $S_{13} \rightarrow S_{14} \equiv$ probable hetero-conciliation of the user's higher-level practical goal).

Thus, the prompt generates cognitive effects analogous to those observed in human interaction. The user interprets their action as communicatively effective, anticipating the likely conciliation of communicative, informational, and practical goals. Although ChatGPT does not intentionally participate in this process, the user's inferential architecture compensates for the lack of reciprocity by projecting goal conciliation as probable. The execution of prompt (3a) therefore functions as an antecedent action that initiates the cycle of assumptions and cognitive effects posited by GCT—even though only one participant conceives of and monitors this inferential chain (Figure 15).

Figure 15. Effects of the User's Action on the User's Intentional Action Plan

Q — Obtain the name of the capital of Spain, User (higher-level practical goal)
P — Request from ChatGPT the name of the capital of Spain, User (lower-level practical goal)
O — Inform ChatGPT that User requests the name of the capital of Spain, User (informational goal)
N — Communicate to ChatGPT that User requests the name of the capital of Spain, User (communicative goal)
M — Issue the prompt "What is the capital of Spain?", User (antecedent action)
M — User issues the prompt "What is the capital of Spain?" (execution of the plan)
N' — User will probably communicate that User requests the name of the capital of Spain, User (communicative goal)
O' — User will probably inform that User requests the name of the capital of Spain, User (informational goal)
P' — User will probably request the name of the capital of Spain, User (lower-level practical goal)
Q' — User will probably obtain the name of the capital of Spain, User (higher-level practical goal)

Source: Authors' own elaboration.

In ChatGPT's case, as noted earlier, the system segments the input into minimal meaning-bearing units (tokens), whose representations are probabilistically related to generate the most likely output sequence. The response-generation process therefore does not involve chains of assumptions or intentional action plans, as it does for human agents; instead, it operates through probabilistic computation based on lexical co-occurrence patterns and contextual weighting.

Although the model's internal processing is purely statistical, its output nevertheless triggers in the user the same relevance-guided interpretive mechanism that governs interpersonal communication. Guided by the principle of relevance, the user interprets the system's utterance as an appropriate response to their request and integrates it into a logical form, thereby deriving the corresponding explicature:

Figure 16. Processing of ChatGPT's "Utterance"

Linguistic Form		The capital of Spain	is	Madrid
		↑	↑	↑
Logical Form		Something	is	something
		↓	↓	↓
Explicature	*ChatGPT asserts that*	the capital of Spain	is	Madrid

Source: Authors' own elaboration.

Assuming the user interprets the model's output as a response to their question, the resulting explicature produces positive cognitive effects within the user's intentional action plan. As expected, the projected conciliation {N–P} of their communicative, informational,

and practical goals is fulfilled. By providing the capital of Spain, ChatGPT enables the user to confirm that they have successfully communicated and conveyed their request (S_{11}–S_{12}) and that their lower-level practical goal—requesting the information (S_{13})—has been achieved.

> S_{15} — ChatGPT asserts that the capital of Spain is Madrid (implicated premise derived from the explicature of its output).
> S_{11} — The user is thereby taken as having communicated their request for the capital of Spain (implicated conclusion by modus ponens, $S_{15}\longrightarrow S_{11}$/strengthening of S_{11} ≡ hetero-conciliation of the user's communicative goal).
> S_{12} — The user is taken as having informed ChatGPT of that request (implicated conclusion by modus ponens, $S_{11}\longrightarrow S_{12}$/strengthening of S_{12} ≡ hetero-conciliation of the user's informational goal).
> S_{13} — The user is taken as having made the request itself (implicated conclusion by modus ponens, $S_{12}\longrightarrow S_{13}$ ≡ hetero-conciliation of the user's lower-level practical goal).

Consequently, the model's response produces in the user a set of cognitive effects equivalent to those observed in active hetero-conciliation. The user recognizes that their practical goal Q has been achieved and that their action plan has been successfully completed.

> S_{14} — The user obtains from ChatGPT the capital of Spain (implicated conclusion by modus ponens, $S_{13}\longrightarrow S_{14}$/strengthening of S_{14} ≡ hetero-conciliation of the user's higher-level practical goal).

On the basis of GCT (Figure 17), the following effects can be identified:

Figure 17. Effects of ChatGPT's Response on the User's Intentional Action Plan

Q — Obtain the name of the capital of Spain, User (higher-level practical goal)
P — Request from ChatGPT the name of the capital of Spain, User (lower-level practical goal)
O — Inform ChatGPT that User requests the name of the capital of Spain, User (informational goal)
N — Communicate to ChatGPT that User requests the name of the capital of Spain, User (communicative goal)
M — Issue the prompt "What is the capital of Spain?", User (antecedent action)
M — User issues the prompt "What is the capital of Spain?" (execution of the plan)
N' — User communicates that User requests the name of the capital of Spain, User (communicative goal)
O' — User informs that User requests the name of the capital of Spain, User (informational goal)
P' — User requests the name of the capital of Spain, User (lower-level practical goal)
Q' — User obtains the name of the capital of Spain, User (higher-level practical goal)

Source: Authors' own elaboration.

Discussion

Interactions with ChatGPT reproduce—with only minimal variation—the pragmatic-cognitive structure observed in the human case. By producing the utterance "I do not know the capital of Spain" (2a), the user explicitly signals a cognitive gap and implicitly issues a request for information. The projected goal is to obtain the capital of Spain, and the enabling abductive hypothesis rests on the expectation that, by marking their ignorance, the system will provide the appropriate response. Within this framework, the utterance operates as an antecedent action in an intentional plan oriented toward the conciliation of that goal.

Response (2b)—"The capital of Spain is Madrid"—confirms this expectation and produces positive cognitive effects analogous to those observed in human interaction. From the user's perspective, active hetero-conciliation occurs: the informational gap is resolved, and the initial hypothesis is confirmed. Practical reasoning unfolds in accordance with GCT, and the processing of the utterance follows Sperber and Wilson's principle of relevance.

The crucial difference, however, is that only the human participant engages in genuine inferential processes. The GPT model does not formulate hypotheses, project goals, or monitor the success of its responses. Its output results from probabilistic computations over token sequences learned from large-scale text corpora, not from cognitive intentionality. The apparent coherence of the response arises from statistical mechanisms that simulate pragmatic inference—without enacting it.

This simulation becomes evident when the system is asked to interpret the dialogue. In explaining its own utterance, the model produces a syntactically well-formed response—"I interpreted your utterance as a statement that you did not know what the capital of Spain was and were seeking that information." The response is linguistically coherent but cognitively empty: the system does not perform the inference it describes; it merely reproduces recurrent explanatory patterns.

This observation suggests that communicative coherence in human–generative-AI interaction rests entirely on human inference. The user interprets the system's response as a sign of understanding and cooperation, whereas what actually occurs is the probabilistic generation of language that merely appears cooperative. This gap between surface appearance and ontological reality defines projected hetero-conciliation: the human mind

attributes agency to a system that only predicts linguistic patterns, thereby sustaining—through cognitive projection—the continuity of the communicative process.

The analysis shows that human–generative-AI interaction reproduces the formal surface of ostensive–inferential communication while shifting its ontological basis. Goal conciliation obtains only for the human participant, who projects onto the system an intentionality it does not possess. Communicative success is therefore functionally projected—a non-reciprocal conciliation sustained by the plasticity of human inferential mechanisms in interaction with a partner that merely simulates agency.

This result suggests that the mechanisms described by RT and GCT remain valid for modeling human cognition in interactions with artificial agents, but it also highlights the need to extend these frameworks to contexts in which communication remains functional without presupposing mind or intentionality in the interlocutor. In such cases, communicative value emerges from the human capacity to project relevance onto systems that merely reflect it at a formal level.

The fundamental difference remains: ChatGPT does not monitor the user's goals, detect instances of conciliation, or confirm hypotheses. When non-conciliation occurs—that is, when the response fails to meet expectations—the user must either reformulate the prompt or abandon the action plan. The entire inferential cycle is unilateral, sustained exclusively by human reasoning.

Even so, the cognitive effects on the human side are functionally equivalent to those observed in genuine conciliation between intentional agents. The cognitive apparatus the user activates—in formulating the abductive hypothesis, producing the utterance, and processing the response—is the same apparatus engaged in human–human interaction. Consequently, GCT remains fully adequate for describing human–generative-AI interaction insofar as it involves user-side processing and the inferential coherence of the action plan.

Final considerations

The analysis presented in this article shows that interactions between humans and generative artificial intelligences reproduce, at the level of their formal surface, the ostensive–inferential structure described by Relevance Theory and the abductive–deductive sequencing outlined by Goal Conciliation Theory. This functional equivalence, however,

depends entirely on human cognitive projection: the individual attributes agency and cooperativeness to the artificial system in order to sustain their own practical reasoning.

It is this projection—rather than any form of mental reciprocity—that maintains communicative coherence. Hetero-conciliation is not eliminated but reconfigured. In the absence of an artificial mind, the human agent assumes full responsibility for sustaining the inferential circuit, heuristically compensating for the system's lack of intentionality.

The notion of projected hetero-conciliation captures this phenomenon. It designates a form of conciliation in which collaboration is inferred rather than generated, yet one that still produces genuine and relevant cognitive effects. This formulation preserves the descriptive power of pragmatic–cognitive theories without resorting to anthropomorphism, while acknowledging both the ontological limits of artificial systems and the interpretive plasticity of human cognition.

Reconceptualizing communication in these terms does not deny the asymmetry between humans and machines; rather, it clarifies how communicative experience persists even when intentionality resides on only one side of the interaction. In this sense, projected hetero-conciliation extends the scope of GCT and RT and invites a rethinking of the cognitive status of communication in an era shaped by generative AI.

The concept also contributes to the renewal of pragmatic–cognitive research by shifting the analytical focus from intentional communication to communication that is functionally interpreted. By recognizing that communicative coherence can arise without mental reciprocity, this perspective allows relevance to be conceived as a relational and dynamic phenomenon—one sustained by heuristic inferences that preserve the continuity of communicative action.

This perspective highlights the adaptability of human pragmatic reasoning in contexts lacking shared intentionality. Projected hetero-conciliation underscores the cognitive flexibility that allows human communication to retain its inferential structure even when mediated by artificial systems devoid of mindedness.

Beyond its theoretical implications, the concept yields methodological and analytical consequences. Demonstrating that relevance and goal conciliation remain operative in the absence of shared intentionality opens avenues for empirically investigating how individuals

sustain their communicative plans when responding to outputs generated through probabilistic prediction.

This approach enables the observation—grounded in real data—of how pragmatic reasoning reorganizes hypotheses, adjusts inferences, and maintains practical coherence when the other pole lacks intentionality. For language studies, this framework extends pragmatic–cognitive analysis to contexts in which agency is merely ascribed, thereby redefining cooperation and relevance in environments mediated by artificial systems.

The example examined in this article constitutes, by design, a controlled test of the model: a minimalist episode marked by relatively stable goals, clearly formulable hypotheses, and low ecological turbulence. This configuration enhances the visibility of GCT's abductive–deductive architecture and allows the analytical isolation of the sequence linking goal projection, hypothesis formulation, and conciliation checking.

However, real-world interactions rarely unfold under such controlled conditions. In everyday communication, goals fluctuate, alternative hypotheses compete, face-related pressures and motivational conflicts reconfigure the action niche, and propositional condensation may temporarily fail. Under such circumstances, the inferential coherence of a hypothesis may prove insufficient to sustain its operability, requiring attention to the ecological dimension of action and to the conditions that enable its continuation.

Recognizing this complexity does not undermine the proposed architecture; rather, it indicates the need to deepen its descriptive–explanatory capacity when applied to less scripted and more turbulent episodes. Questions thus remain regarding the emergence, oscillation, and breakdown of antefactual hypotheses in unstable ecologies—questions that point toward a research agenda still in need of systematic development.

References

NIERO, G. **"Você pode passar o sal?"**: polidez, relevância e heteroconciliação de metas. 2022. 175 f. Tese (Programa de Pós-graduação em Ciências da Linguagem) — Universidade do Sul de Santa Catarina, UNISUL, Tubarão, 2022.

PEDROSO, F. M. **Conciliação de metas e interação entre humanos e inteligências artificiais generativas**: potencialidades descritivo-explanatórias. 2025. 85 f. Dissertação (Mestrado em Ciências da Linguagem) — Universidade do Sul de Santa Catarina, 2025.

RAUEN, F. J. For a goal conciliation theory: ante-factual abductive hypotheses and proactive modelling. **Linguagem em (Dis)curso**, v. 14, n. 3, p. 595-615, set./dez. 2014. DOI: https://doi.org/10.1590/1982-4017-140309-0914.

RAUEN, F. J. Intenção e Conciliação de Metas. **Percursos Linguísticos**, v. 10, p. 24-48, 2020. DOI: https://doi.org/10.47456/pl.v10i26.32867.

RAUEN, F. J. Por uma modelação abdutivo-dedutiva de interações comunicativas. *In*: TENUTA, A. M.; COELHO, S. M. (Org.). **Uma abordagem cognitiva da linguagem** [livro eletrônico]: perspectivas teóricas e descritivas. Belo Horizonte: FALE/UFMG, 2018. p. 13-29.

SPERBER, D.; WILSON, D. **Relevance**: communication and cognition. 2nd. ed. Oxford: Blackwell, 1995.

Sobre os autores

Francine Mates Pedroso

Orcid: https://orcid.org/0009-0008-7556-6136

Francine Pedroso holds a Master's degree in Language Sciences and is currently a PhD student in the Graduate Program in Language Sciences at the University of Southern Santa Catarina (UniSul), supported by a PROSUC/CAPES scholarship. Her interdisciplinary research focuses on Relevance Theory, Goal Conciliation Theory, distributed cognition, human–machine interaction, and linguistic approaches to artificial intelligence.

Fábio José Rauen

Orcid: https://orcid.org/0000-0002-1096-7253

Fábio Rauen holds a PhD in Linguistics and is a professor and coordinator of the Graduate Program in Language Sciences at the University of Southern Santa Catarina (UniSul), where he also serves as editor of the journal Linguagem em (Discurso). A researcher at the Ânima Institute, he is known for his work in cognitive pragmatics and for proposing Goal Conciliation Theory. He is the author of reference books in academic writing and research methodology.