The Gricean Maxims in NLP - A Survey

Lea Krause
Vrije Universiteit Amsterdam
l.krause@vu.nl

Piek Vossen
Vrije Universiteit Amsterdam
p.t.j.m.vossen@vu.nl

Abstract

In this paper, we provide an in-depth review of how the Gricean maxims have been used to develop and evaluate Natural Language Processing (NLP) systems. Originating from the domain of pragmatics, the Gricean maxims are foundational principles aimed at optimising communicative effectiveness, encompassing the maxims of Quantity, Quality, Relation, and Manner. We explore how these principles are operationalised within NLP through the development of data sets, benchmarks, qualitative evaluation and the formulation of tasks such as Data-to-text, Referring Expressions, Conversational Agents, and Reasoning with a specific focus on Natural Language Generation (NLG). We further present current works on the integration of these maxims in the design and assessment of Large Language Models (LLMs), highlighting their potential influence on enhancing model performance and interaction capabilities. Additionally, this paper identifies and discusses relevant challenges and opportunities, with a special emphasis on the cultural adaptation and contextual applicability of the Gricean maxims. While they have been widely used in different NLP applications, we present the first comprehensive survey of the Gricean maxims' impact.

1 Introduction

Capturing the full nuance of human language requires more than understanding its structure; it necessitates an intricate comprehension of context. This understanding goes beyond the words themselves to grasp the intentions, implications, and subtleties embedded in communication (Wittgenstein, 1953; Grice, 1975; Levinson, 2000).

In order to build NLP systems that are able to use language beyond just its literal content, they need to incorporate pragmatic capabilities (Hovy, 1987, 1990; Hovy and Yang, 2021; Pritzkau et al., 2023; Seals and Shalin, 2023). A central idea in pragmatics are the Gricean maxims, a set of cooperative principles proposed by philosopher Grice (1975). These maxims are descriptions of effective human communication strategies, capturing the implicit expectations and norms that govern human interaction and thereby offering a theoretical framework that has profound implications for the development of NLP technologies. As NLP systems, particularly LLMs, strive to achieve more human-like understanding and generation of text, the consideration of these pragmatic principles becomes crucial (Jacquet et al., 2019b; Kasirzadeh and Gabriel, 2023; Alexandris, 2024). They not only aid in improving the interpretative capabilities of these systems but also enhance their ability to generate coherent, contextually appropriate responses.

The Gricean maxims consist of four primary directives that guide conversational cooperation. Each maxim addresses a different aspect of communication, providing a guideline for what makes a conversation effective and meaningful. These maxims are:

Maxim	Description
Quantity	Make your contribution as informative as necessary, without providing excessive information.
Quality	Ensure your contribution is true and based on evidence.
Relation	Your contribution should be relevant to the conversation.
Manner	Your contribution should be clear, concise, and orderly, avoiding ambiguity and obscurity.

Table 1: Gricean maxims and their descriptions

Maxim of Quantity stresses the importance of providing an appropriate amount of information. Too little information can leave the listener confused or in need of clarification, while too much can overwhelm or distract. In summarisation tasks,

Proceedings of the 17th International Natural Language Generation Conference, pages 470–485
September 23–27, 2024.

this maxim guides systems to include all critical data without including superfluous detail, ensuring summaries are both comprehensive and focused.

Maxim of Quality deals with the truthfulness and reliability of the communicated message. It discourages the sharing of falsehoods or unfounded assertions. In the context of data-to-text generation, using this maxim can ensure that texts are based on accurate data and that any predictive or inferential statements have a solid basis in the available information.

Maxim of Relation, also known as relevance, mandates that contributions be pertinent to the current topic of discourse. This principle is particularly relevant in question-answering systems and conversational agents, where responses must directly address the user's queries or comments to maintain a coherent and contextually appropriate dialogue.

Maxim of Manner emphasises the way information is presented, advocating for clarity, brevity, and orderliness. This maxim can help in generating user-friendly texts, avoiding jargon, overly complex structures, or ambiguous phrasing that could hinder comprehension. It supports the design of systems that produce outputs easy for the end-user to understand and act upon.

Collectively, these maxims provide a valuable heuristic for designing and evaluating NLP systems, from chatbots and conversational agents to summarisation and translation tools. They ensure that automated systems not only generate human-like text but also engage in human-like conversation dynamics, ultimately aiming for natural, efficient, and effective communication.

This paper systematically examines the influence of Gricean maxims across various facets of NLP. We explore:

Data and Benchmarks: The construction and evaluation of datasets and benchmarks grounded in pragmatic principles.

Tasks: Covering NLP tasks in NLG such as data-to-text, summarisation, translation, referring expressions, and related fields such as NLU and conversational AI, we discuss how the Gricean maxims inform works in these areas.

LLMs: The application and impact of Gricean principles in the development and assessment of current large language models.

Criticisms and Future Work: We highlight shortcomings and potential for future research, particularly focusing on the cultural adaptation of Gricean maxims, which could inform more nuanced and globally applicable NLP systems.

With the present survey, our aim is to underscore the potential of the Gricean maxims in enhancing the communicative and interpretative faculties of NLP systems, making them more effective and context-aware in their language use. By giving the first comprehensive overview of existing work, we hope to enable future research in this area.

2 Methodology

We compile our list of papers through an exhaustive keyword search on Google Scholar and the ACL Anthology database. We combined keywords for the concepts (`Gricean maxims`, `Cooperative principles`, `Pragmatic principles`) with keywords for disciplines (`NLP`, `NLG`, `Conversational AI`) in a two-dimensional matrix. After manually filtering out papers that only mention the Gricean maxims in their related work or introduction and additions through mentioned related work, we identified 78 relevant papers published between 1990 and 2024. For an overview of all works surveyed, see Figure 1. For a division into the covered maxims see Appendix 2.

3 Data and Benchmarks

In this section, we show recent advancements in the creation of datasets and benchmarks aimed at evaluating and enhancing the pragmatic reasoning capabilities of NLP systems, particularly LLMs.

GRICE Dataset Zheng et al. (2021) present the GRICE dataset, a grammar-based dialogue dataset designed to incorporate implicature into pragmatic reasoning within conversations. The dataset aims to bridge the gap in modern open-ended dialogue systems that struggle with understanding the intended meaning beyond the literal statements. GRICE also addresses other crucial aspects of dialogue modelling, such as coreference, ensuring temporal consistency and intricate implicatures within each dialogue context. The dataset introduces two tasks: implicature recovery and pragmatic reasoning in conversation. Experiments reveal a significant gap between the performance of baseline methods (which claim pragmatics reasoning capabilities) and human performance. Incorporating

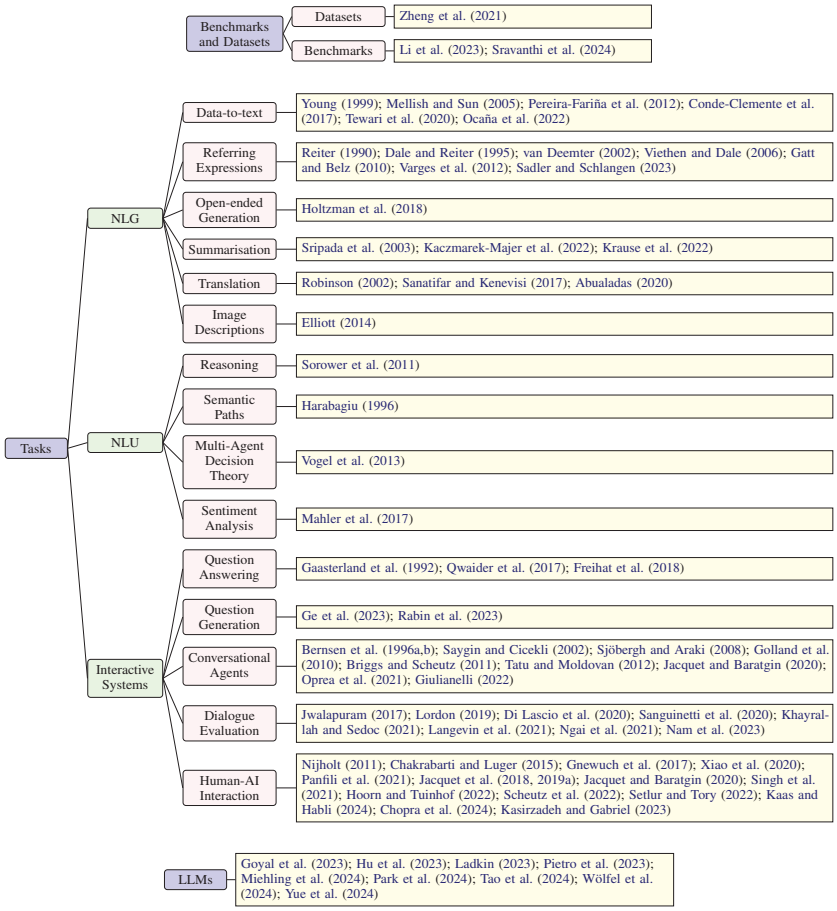

Figure 1: Overview of papers surveyed. The classification of papers is not strictly exclusive, as work from both Interactive Systems and LLMs overlaps with tasks in NLG and NLU.

a module for explicit implicature reasoning shows to significantly improve conversational reasoning performance.

DiPlomat Benchmark Li et al. (2023) introduce the DiPlomat benchmark to enhance conversational agents' understanding and reasoning with nuanced and ambiguous language. It targets three key areas: situational context reasoning, open-world knowledge acquisition, and figurative language understanding. The benchmark includes a human-annotated dataset of 4,177 multi-turn dialogues with a 48,900-word vocabulary. It features tasks such as Pragmatic Reasoning and Identification and Conversational Question Answering, plus a zero-shot natural language inference task emphasising context's role in pragmatic reasoning. Results highlight current LLMs' limitations in this area.

Pragmatics Understanding Benchmark (PUB) Sravanthi et al. (2024) release the Pragmatics Understanding Benchmark to illustrate LLMs' challenges in grasping pragmatic aspects of language, despite their proficiency in understanding semantics. PUB encompasses fourteen tasks across four pragmatic phenomena: Implicature, Presupposition, Reference, and Deixis. With a total of 28k data points, including 6.1k created by the authors and the rest adapted from existing datasets, PUB serves as a comprehensive testbed for evaluating LLMs' pragmatic reasoning abilities. The benchmark's findings indicate that while fine-tuning for instruction-following and chat improves smaller models' pragmatics capabilities, larger models show comparable performance between their base and chat-adapted versions. However, a notable gap exists between the models' capabilities and human performance, with models displaying variability in proficiency across different tasks and complexity levels within the same dataset.

4 Tasks

In NLP, Gricean maxims are widely applied in various tasks, particularly in NLG. These maxims are relevant to NLG (4.1) because they help generate text that adheres to human conversational norms, making interactions more intuitive and effective. By following these principles, NLG systems produce responses that are clear, relevant, and contextually appropriate, thereby enhancing the naturalness and coherence of the generated language. In NLU (4.2), Gricean maxims can enhance some interpretive tasks, such as reasoning, decision-making, and sentiment analysis, by improving the processing of language in a way that mirrors human understanding. These cooperative principles are thus also applicable in Interactive Systems (4.3), like Question-Answering or Conversational Agents, which integrate both NLG and NLU to create seamless and coherent interactions. We review a wide range of works, showing the broad applicability of Grice's cooperative principles.

4.1 NLG

Data-to-text In the domain of data-to-text generation, adherence to Grice's maxims ensures the production of linguistic reports that are both accurate and user-oriented. An early approach by Young (1999) focuses on generating textual descriptions of complex activities, employing Grice's maxim of Quantity to produce cooperative plan descriptions that are concise yet informative. This approach uses a computational model of the hearer's plan reasoning capabilities to select the most appropriate plan descriptions, emphasising the collaborative nature of communication. The work by Mellish and Sun (2005) on Natural Language Directed Inference deals with content determination: selecting relevant material for inclusion in the system's final natural language output. They describe their desiderata as potential cases of the Gricean maxims, akin to the approach taken by Sripada et al. (2003) for summarisation. Pereira-Fariña et al. (2012) and Conde-Clemente et al. (2017) assess the quality of linguistic reports generated from vehicle simulator data and big data respectively, applying the Gricean maxims as evaluative criteria. These studies highlight the complexities of ensuring quality in linguistic reports, showing that adherence to Grice's maxims can address issues such as scalability, efficient processing, and the relevance of information, thereby enhancing the intuitiveness and effectiveness of the generated reports. Tewari et al. (2020) explore the Quantity maxim's role in informativeness, particularly in navigation instructions. They propose metrics for evaluating syntactic cohesion and informativeness, finding that simple syntactic measures align well with human judgements of instruction quality.

Referring Expressions Reiter (1990) provides a foundational interpretation of the Gricean maxims for generating referring expressions. They emphasise the need for these expressions to be brief, avoid unnecessary elements, and use preferred lexical classes to prevent false conversational implicatures. They formalise these principles into three preference rules: Local Brevity, No Unnecessary Components, and Lexical Preference, and integrate them into a polynomial-time algorithm for generating accurate referring expressions. Dale and Reiter (1995) build on this work, examining various computational interpretations of the Gricean maxims to generate definite noun phrases that similarly identify intended referents without causing false implicatures. They conclude that the simplest and fastest interpretation often aligns best with human conversational behaviour and present the efficient and adaptable Incremental Algorithm for this purpose. Proving its adaptability, it was for example implemented for the automatic generation of medical reports (Varges et al., 2012) and used in the creation of the diagnostic dataset Pento-DIARef (Sadler and Schlangen, 2023). Further extending the algorithm, van Deemter (2002) incorporates Boolean logic to enhance informativeness and relevance, ensuring the generated expressions are both clear and contextually appropriate. To evaluate the performance of existing algorithms, Viethen and Dale (2006) present a dataset of human-produced referring expressions, noting significant differences between human and algorithm-generated expressions. For a shared task, Gatt and Belz (2010) evaluate REG systems by applying theoretically motivated criteria based on the Gricean Maxim of Quantity. They measure the minimality of attribute sets, ensuring that descriptions include no more information than required for identification.

Open-ended Generation An initiative to create a more powerful generative model builds upon the foundation of an RNN language model, incorporating discriminative models inspired by Grice's maxims (Holtzman et al., 2018). This setup aims to

produce language that is coherent, informative, and contextually relevant, marking a departure from generic responses. Evaluations suggest that language generated by this model is preferred by users over competitive baselines, offering improvements in coherence, style, and information content.

Summarisation The Gricean maxims, particularly those of Quantity, Relation, and Manner, find significant application in the field of text summarisation, guiding systems towards generating concise and contextually relevant summaries. Sripada et al. (2003) highlight an operational weather-forecast generator that selects trends and patterns, converting these into linguistic expressions for textual summaries. This process, rooted in Gricean maxims, ensures communication with users is clear, informative, and pertinent, showcasing the maxims' role in enhancing data-to-text communication. Krause et al. (2022) focus on list verbalisation in Knowledge Graph QA systems, addressing the challenge of summarising too many potential answers to open questions. Their approach, informed by Gricean maxims, employs graph-based and language model-based measures to rank answers, emphasising the need to balance content that is both popular and contextually appropriate.

Translation In translation studies, the Gricean maxims are utilised as analytical tools to navigate the pragmatic complexities involved in transferring meaning across languages. Robinson (2002) discusses the application of these maxims in translation, emphasising the translator's challenge to preserve or adapt the original author's violations of these maxims to maintain the intended implicatures in the target text. This approach underscores the role of pragmatic implicature for translators to effectively communicate the original message to a new audience. Sanatifar and Kenevisi (2017) address the cultural nuances of applying Grice's maxims in translation, suggesting a reformulation within a framework of faithfulness to make them more adaptable to the diverse needs of translation. Their analysis of examples from translations showcases the potential adjustments needed to align these maxims with the specific requirements of translation tasks. For fiction translation, Abualadas (2020) explore the application of Grice's maxims in the Arabic translations of "Animal Farm," investigating the communicative principles underlying character-to-character, narrator-to-reader, and translator-to-reader interactions. The study reveals a higher level of explicitness and informativeness in the translations, indicating the translators' efforts to adhere to conversational maxims during the mediation process, albeit with a noted increase in explicitness that may affect reader engagement and the persuasive power of the text.

Image Descriptions Elliott (2014) provide an overview of the image description literature through the lens of Grice's maxims. They critique current models for focusing mainly on semantic correctness and relevance, neglecting the maxim of Quantity, which results in overly detailed descriptions. They stress the need for evaluation models that balance all maxims, noting that as computer vision accuracy improves, the distinction between relevant, quality descriptions and those of adequate quantity becomes crucial, a nuance often missed in current human judgements but adhered to in gold-standard crowdsourced descriptions.

Human Evaluation Across NLG tasks the maxims have also been utilised as guidance for human evaluation of generated language (van der Lee et al., 2021), e.g. reports or summaries (Ocaña et al., 2022; Kaczmarek-Majer et al., 2022). Most recently, Google's LaMDA (Thoppilan et al., 2022) system's metrics (Sensibleness, Specificity and Interestingness) for human evaluation have been mapped to the Gricean maxims (Wahlster, 2023). See also *Dialogue Evaluation* in Section 4.3.

4.2 NLU

Reasoning Work on inverting Grice's maxims to learn rules from natural language texts (Sorower et al., 2011) highlights a novel approach to extracting domain knowledge from concise information sources like news articles. This method models the probability of facts being mentioned, leveraging the understanding that texts often contain just enough information for readers to infer the missing pieces based on shared knowledge. By formalising the maxims of truthfulness and conciseness, this approach successfully infers more information from texts than standard methods, illustrating the applicability of Grice's maxims in learning from incomplete data.

Sentiment Analysis In exploring strategies to challenge sentiment analysis systems, Mahler et al. (2017) employed linguistic manipulations based on

Gricean principles. By editing test data to create instances where conversational maxims are flouted, the study assessed the systems' abilities to interpret the underlying sentiment correctly. This approach revealed significant challenges for NLP systems, especially when dealing with semantic and pragmatic manipulations that subtly convey sentiment through the violation of Grice's maxims.

Multi-Agent Decision Theory Research into multi-agent decision-making demonstrates how the cooperative principle and Grice's maxims of Relevance, Quality, and Quantity naturally emerge from decision processes involving multiple agents (Vogel et al., 2013). Using a decentralised decision-making model, the study shows that agents' reasoning about each other's beliefs and intentions—aligned with Gricean communicative behaviour—significantly improves task performance.

Semantic Paths A proposal for using Gricean maxims to validate semantic paths in knowledge bases underscores the potential for these principles to ensure coherence and relevance in information retrieval (Harabagiu, 1996). This approach posits that Gricean maxims can serve as a filter for irrelevant information, facilitating more effective and contextually appropriate responses from knowledge-based systems.

4.3 Interactive Systems

Combining aspects and tasks from both NLG and NLU, Interactive Systems can also be developed and evaluated according to the Gricean maxims, as the cooperative principles can guide effective communication between user and system.

Question-Answering In question-answering systems, Gricean maxims serve as guiding principles to enhance the interaction between users and databases or information systems. Early work by Gaasterland et al. (1992) highlights the importance of cooperative behaviour in these systems, advocating for responses that go beyond direct answers to include extra or alternative information that aligns with the users' needs and expectations. This approach, rooted in the maxims, aims to make these systems more user-friendly and efficient in delivering relevant information. Following this foundational work, Qwaider et al. (2017) apply Gricean principles to rank answers in community question-answering forums. They use semantic similarity and polarity terms to evaluate responses based on the maxims of Quantity, Relation, and Manner, aiming to identify the most informative and contextually appropriate answers. Freihat et al. (2018) explores the application of Grice's maxims from an engineering perspective, focusing on the extensional relevancy of answers to rank them according to their informativeness.

Question Generation Gricean-inspired evaluation metrics are proposed for generating follow-up questions in conversational surveys (Ge et al., 2023), leading to more dynamic and personalised experiences. In an educational setting, Rabin et al. (2023) propose a model that generates gap-focused questions (GFQs) to facilitate effective dialogue. They base their discourse desiderata on the maxims of Relevance, Quantity, and Manner to ensure the answerability of the question, and that while the answers should not yet be in the common ground, all the information used in the question should be.

Conversational Agents The development of conversational agents has long explored aligning with Gricean maxims to ensure natural and effective user-agent communication. Bernsen et al. (1996a) explore how new maxims formulated for human-bot dialogues relate to Gricean principles, emphasising the preservation of the Quantity maxim to ensure unambiguous and contributing responses in conversations. Further refining these ideas, Bernsen et al. (1996b) present a set of principles for cooperative spoken human-machine dialogue, developed through user testing and comparisons with human-human dialogue theory. These principles extend Grice's Cooperative Principle, addressing specific aspects of dialogue not covered by the original maxims and offering a practical framework for designing and evaluating spoken dialogue systems. The application of Gricean maxims in designing conversational agents has been further explored by Saygin and Cicekli (2002), who provide a pragmatic analysis of human-computer conversations. They examine how computers' violations of the maxims affect their ability to imitate human conversational behaviour, highlighting the challenges and requirements for conversational agents to successfully cooperate within human communication frameworks. In Golland et al. (2010) a game-theoretic model where a rational speaker generates utterances by considering the listener's perspective according to the Maxim of Manner significantly outperforms a baseline reflex speaker in generat-

ing spatial descriptions. In the context of mental modelling, Briggs and Scheutz (2011) introduce an algorithm that integrates belief revision and expression, enabling robots to monitor and update the beliefs of their conversation partners while adhering to Gricean maxims of language use. Similarly, Giulianelli (2022) propose the development of NLG systems that learn pragmatic production decisions through experience, by evaluating goals, costs, and utility in a human-like fashion, and show how their framework and cost model map to the Gricean maxims. Jacquet and Baratgin (2020) propose a chatbot model aimed at enhancing the pragmatic aspects of language processing, stressing the importance of distinguishing between sentence processing and information processing, to generate responses that address the user's informational needs and situational context simultaneously.

Implied meanings, which are not directly stated but understood from context, present a significant challenge for conversational agents due to their reliance on subtle cues and contextual knowledge. For example, Gricean maxims have been applied to humour generation with moderate success (Sjöbergh and Araki, 2008). Tatu and Moldovan (2012) explore the extraction of conversational implicatures, advancing the ability of conversational agents to discern and convey implied meanings within dialogues. Their work enhances the agents' interpretative layer, allowing for a deeper understanding of the subtleties present in human conversations. Sarcasm, as an extreme form of implied meaning, introduces additional complexity. With Chandler, Oprea et al. (2021) introduce a system adept at sarcastic response generation, which moves away from the traditional understanding of sarcasm in light of Grice's quality maxim and instead focuses on the crucial role of intention behind utterances.

Dialogue Evaluation Evaluating conversational agents for their adherence to Gricean maxims provides insights into their effectiveness and user satisfaction. Many works propose frameworks where human raters assess dialogues based on Gricean categories (Jwalapuram, 2017; Lordon, 2019; Langevin et al., 2021; Ngai et al., 2021; Nam et al., 2023). Additionally, Sanguinetti et al. (2020) and Di Lascio et al. (2020) cluster error types for tagging into a coarse-grained taxonomy inspired by the maxims. Through their metric called Relative Utterance Quantity (RUQ), Khayrallah and Sedoc (2021) assess a model's preference for generic "I don't know" responses even when more informative responses are available, classifying them as a failure to adhere to the Maxim of Quantity.

Human-AI Interaction Human-AI interactions provide a rich area for applying and testing Gricean maxims, offering insights into how these principles influence user satisfaction and system performance in real-world settings. In costumer service, Chakrabarti and Luger (2015) and Gnewuch et al. (2017) focus on designing conversational agents that improve service quality by understanding the context and intent behind conversations by drawing on the cooperative principle and social response theory, they propose design principles for agents that can engage users in a more meaningful and contextually relevant manner. Xiao et al. (2020) explore the effectiveness of chatbots in surveys, finding that adherence to Gricean maxims results in higher engagement and response quality. Similarly, Panfili et al. (2021)'s study revealed that violations of Grice's maxims in interactions with Alexa led to user frustration, with Relevance violations being particularly aggravating. Building on this, Jacquet et al. (2018) and subsequent studies by overlapping authors in 2019a and 2020, further explore the cognitive dimensions of human-AI communication. They examine how deviations from Gricean principles impact response times and cognitive load, demonstrating that violations, especially of the Relation and Quantity maxims, can significantly burden the interaction process. Their work highlights the cognitive cost of processing information when conversational norms are not met, suggesting that AI systems should minimise these violations to facilitate smoother and more natural dialogues. Focusing on the Maxim of Quantity, Singh et al. (2021) present a mechanism for robot teams to verbalise and explain their actions and intentions to improve human understanding, showing that this approach, implemented on three Pepper robots (Pandey et al., 2018), results in the greatest comprehension compared to other methods. This sort of explanation transparency is likewise stressed in (Scheutz et al., 2022), especially when rejecting human commands. Unlike Singh et al. (2021), they incorporate all Gricean maxims in their definition of transparency. In another framework, the Maxim of Quality is used in an intentional operator to keep an interaction from failing if the agent encounters uncertainty about conflicts in a user's statements

and its ontology (Hoorn and Tuinhof, 2022). Setlur and Tory (2022) study how Gricean maxims can guide the design of chatbot interfaces for data exploration. By employing cooperative principles, they aim to create chatbots that better support users' information-seeking behaviours, adapted to specific modalities like text and voice. Their Wizard of Oz studies (Dahlbäck et al., 1993) reveal user preferences for intent interpretation and highlight the need for chatbot design to adapt based on interface affordances, ensuring that interactions are both informative and contextually appropriate. The maxims have also been used to structure effective responses when communicating about AI safety to diverse stakeholders (Kaas and Habli, 2024) or about bugs to developers (Chopra et al., 2024).

Kasirzadeh and Gabriel (2023) explore the aligning of conversational agents with Gricean maxims more critically, emphasizing the need for context-specific adaptation. They argue that while Gricean maxims offer a foundational framework for designing aligned conversational agents, the application of these principles is not straightforward due to contextual variations and propose a principle-based approach, highlighting the importance of understanding how these maxims operate in different domains. Similarly, Goodman and Frank (2016) suggest the use of the Rational Speech Act model, which replaces Grice's maxims with a utility-theoretic cooperative principle that reflects the communicative and social priorities of real-world agents. Lastly, sometimes people will purposefully not follow cooperative principles. Hence, in conversational settings with a virtual agent or social robot, it is beneficial for the artificial partner to accept that its human counterpart might not follow the Gricean principles and adapt accordingly (Nijholt, 2011).

5 LLMs

With the widespread use of LLMs, expectations are emerging for them to have pragmatic abilities: to interpret and generate language in context. In the following, we look at recent approaches that use Gricean maxims to evaluate and potentially improve these capabilities.

Hu et al. (2023) perform an in-depth evaluation comparing the performance of LLMs with humans across a spectrum of pragmatic phenomena. Their research reveals that top-tier models match humans in terms of accuracy and error tendencies, showing a preference for literal over heuristic interpretations. However, challenges arise with scenarios that demand an understanding of violated social norms.

Similar gaps are found by Pietro et al. (2023) when analysing ChatGPT's grasp on pragmatics, identifying its proficiency across various domains but pinpointing deficits in understanding humour, metaphors, and adhering to the quantity maxim. Tao et al. (2024) corroborate these findings with a naturalness metric that is based on the cooperative principles and the model again most frequently violating the Maxim of Quantity. Investigating the reverse, Yue et al. (2024) studied if LLMs can spot maxim violations and implicatures. They find that while the performance of LLMs did not significantly vary with respect to different conversational maxims, variability existed in the performance among models. Miehling et al. (2024) propose an augmented set of conversational maxims to evaluate and guide interactions between humans and LLM-driven conversational agents, adding maxims for Benevolence (to avoid harm) and Transparency (admitting limitations).

Gricean maxims are also used in critiques of the application of LLMs like ChatGPT in legal environments (Ladkin, 2023), focusing on its tendency to produce unverified content, termed "r-lying." This critique leverages Grice's Quality maxim to question the reliability and accuracy of responses generated by LLMs, underlining the imperative for technological advancements to mitigate these issues. In entity description generation, Goyal et al. (2023) adapt the maxims of Quality and Relation into factuality and congruity. Their evaluation paradigm disentangles factual errors (nonfactual descriptions) from contextual errors (incongruous descriptions). They find that models struggle with accurate descriptions of less familiar entities, raising concerns about the trustworthiness of language models, as these errors are harder for human readers to detect.

In multilingual contexts, Park et al. (2024) broaden the scope of assessing LLMs' pragmatic skills to include Korean, utilising diverse question formats to test narrative response capabilities. Their study demonstrates GPT-4's strong performance, while cautioning against prompting methods that skew towards literal interpretations, thus limiting pragmatic inference. In the educational domain in German, AI-driven pedagogical agents are evaluated by applying Gricean principles extended with a Trust maxim (Wölfel et al., 2024). The findings emphasise trust as a crucial factor in

the educational efficacy of chatbots, suggesting that fidelity to Gricean norms can significantly impact the utility and dependability of conversational AI in learning contexts.

These investigations collectively underscore the ongoing effort to give LLMs a deeper understanding of pragmatic nuance. While significant progress has been made, the reviewed works highlight the particular relevance of the maxim of Quantity in addressing overgeneration issues in LLMs (Pietro et al., 2023; Tao et al., 2024) and the application of the maxim of Quality in improving their expression of uncertainty (Hoorn and Tuinhof, 2022). Achieving full pragmatic alignment remains a challenge, pointing to future research directions that could bridge the gaps in current capabilities.

6 Criticisms and Future Work

Pragmatic Criticisms The Gricean maxims should be interpreted within the broader context of pragmatic theory, acknowledging that while foundational, they face criticism and alternatives.

As argued by Davies (2000), there is a need to distinguish between the colloquial use of "cooperation" and the use intended by Grice, a distinction he terms "cooperation drift." Similarly, Chen and van Deemter (2023) emphasise the need for explicit definitions of over- and under-specifications in referring expressions, noting that these are often loosely aligned with Gricean principles without clarifying "required" actually means.

Neo-Griceans (Horn, 1972; Atlas and Levinson, 1981) simplify Grice's maxims into two principles: the Q-principle, which encourages providing sufficient information while avoiding unnecessary details, and the I-principle, which emphasizes clarity and informativeness. This approach aims to create a more unified and manageable framework for understanding conversational implicatures

Additionally, some scholars argue that Grice's maxims are vague and oversimplify communication complexities (Frederking, 2004). Others question their universal applicability, noting real-world deviations (Levinson, 2000), or the dynamic negotiation of meaning that sometimes breaks these maxims to achieve understanding (Clark, 1996). Power dynamics and politeness strategies, which also influence conversations, are insufficiently addressed by Grice's framework (Leech, 1983; Brown and Levinson, 1987).

Cultural Adaptation As mentioned by Hovy and Yang (2021) culture and language are fused, thereby making a language analysis without looking at the social and cultural aspects of it limited in its insights. This also holds for the Gricean maxims. As Danziger (2010) documents, while the maxims were intended as universal, certain cultural settings might interpret the maxims differently, indicating a need for cultural adaptation of these principles. A promising way to deal with this is participatory design, where stakeholders affected by AI systems should participate in their design (Delgado et al., 2023). An example of an application with relevance to the Gricean maxims is the study by Medhi Thies et al. (2017) who explored chatbot preferences in an exploratory Wizard-of-Oz study among young, urban Indians. Machali (2012), Olaniyi and Oyinbo (2021), and Kamal and Mhamed (2023) contribute to the discourse by examining the structure of Grice's Maxims within the Indonesian, Nigerian and Moroccan cultural contexts, respectively. Their findings highlight the influence of societal expectations, politeness strategies, and specific linguistic characteristics on conversational implicatures, suggesting that the maxims may require re-formulation or adaptation to align with diverse context dependent cultural norms.

7 Conclusion

The application of Gricean maxims in NLP reflects a consistent effort to address the complexities of human communication, spanning from the foundational stages of the field to current advancements. This survey is the first comprehensive review of how these maxims have informed the development and evaluation of NLP systems across a range of tasks, highlighting progress in making systems more pragmatically aligned with human conversation while also pointing out the existing challenges, especially regarding cultural variations and conversational norms. While some papers focus on specific subsets of the maxims, others extend them to capture a broader spectrum of communicative nuances or reinterpret them for their use-case. Surveyed work suggests that moving forward, the NLP field can benefit from a more focused integration of pragmatic and cultural considerations, aiming to produce conversational agents that better reflect the intricacies of human communication.

Acknowledgments

This research was funded by the Vrije Universiteit Amsterdam and the Netherlands Organisation for Scientific Research (NWO) through the *Hybrid Intelligence Centre* via the Zwaartekracht grant (024.004.022), and the *Spinoza* grant (SPI 63-260) awarded to Piek Vossen.

Limitations

One significant criticism is the potential cultural specificity of Gricean maxims. Research has suggested that the assumptions underpinning these maxims may not hold universally across different languages and cultural communication norms. This indicates a limitation in applying Gricean principles as a one-size-fits-all framework for conversational agents intended for a global audience. It raises the question of whether these maxims can fully capture the nuances of non-Western communication styles or the subtleties of multilingual discourse. This survey is impacted by this, as the majority of works surveyed are done in English speaking or Western contexts and might not hold when generalised to other cultural contexts.

Moreover, the Gricean framework primarily focuses on the ideal cooperative conversation without accounting for the complexity of real-world interactions that may involve conflict, competition, or deception. This gap suggests the need for integrating additional pragmatic theories that can accommodate a wider range of communicative intentions and strategies beyond cooperation.

Furthermore, the operationalisation of Gricean maxims in NLP often relies on simplified or binary interpretations of these principles, which may not fully encapsulate their intended scope or the dynamic nature of pragmatics. This simplification can lead to challenges in addressing the subtleties of conversational implicature or the fluidity of context in automated language processing tasks.

In conclusion, while the application of Gricean maxims offers valuable insights into the pragmatics of language use in computational contexts, it is imperative to recognise their limitations and the importance of exploring a broader spectrum of pragmatic theories.

Supplementary Materials Availability Statement: For reproducibility the keyword combinations mentioned in 2, should be searched on https://scholar.google.com and https://aclanthology.org. The final search was done in May 2024.

References

Othman Ahmad Abualadas. 2020. Conversational maxims in fiction translation: New insights into cooperation, characterization, and style. *Indonesian Journal of Applied Linguistics*, 9(3):637–645.

Christina Alexandris. 2024. GenAI and Socially Responsible AI in Natural Language Processing Applications: A Linguistic Perspective. *Proceedings of the AAAI Symposium Series*, 3(1):330–337.

Jay David Atlas and Stephen C Levinson. 1981. It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In *Radical pragmatics*, pages 1–62. Academic Press.

Niels Ole Bernsen, Hans Dybkjær, and Laila Dybkjær. 1996a. Cooperativity in human-machine and human-human spoken dialogue. *Discourse Processes*, 21(2):213–236.

N.O. Bernsen, H. Dybkjaer, and L. Dybkjaer. 1996b. Principles for the design of cooperative spoken human-machine dialogue. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*, volume 2, pages 729–732 vol.2.

Gordon Briggs and Matthias Scheutz. 2011. Facilitating Mental Modeling in Collaborative Human-Robot Interaction through Adverbial Cues. In *Proceedings of the SIGDIAL 2011 Conference*, pages 239–247, Portland, Oregon. Association for Computational Linguistics.

Penelope Brown and Stephen C Levinson. 1987. *Politeness: Some universals in language usage*, volume 4. Cambridge university press.

Chayan Chakrabarti and George F. Luger. 2015. Artificial conversations for customer service chatter bots. *Expert Systems with Applications: An International Journal*, 42(20):6878–6897.

Guanyi Chen and Kees van Deemter. 2023. Varieties of specification: Redefining over- and under-specification. *JOURNAL OF PRAGMATICS*, 216:21–42.

Bhavya Chopra, Yasharth Bajpai, Param Biyani, Gustavo Soares, Arjun Radhakrishna, Chris Parnin, and Sumit Gulwani. 2024. Exploring Interaction Patterns for Debugging: Enhancing Conversational Capabilities of AI-assistants. *Preprint*, arxiv:2402.06229.

Herbert H Clark. 1996. *Using language*. Cambridge university press.

Patricia Conde-Clemente, Gracian Trivino, and Jose M. Alonso. 2017. Generating automatic linguistic descriptions with big data. *Information Sciences*, 380:12–30.

Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of oz studies: why and how. In *Proceedings of the 1st international conference on Intelligent user interfaces*, pages 193–200.

Robert Dale and Ehud Reiter. 1995. Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 19(2):233–263.

Eve Danziger. 2010. On trying and lying: Cultural configurations of Grice's Maxim of Quality. *Intercultural Pragmatics*, 7(2):199–219.

Bethan Davies. 2000. Grice's cooperative principle: Getting the meaning across. *Leeds Working Papers in Linguistics and Phonetics*, 8.

Fernando Delgado, Stephen Yang, Michael Madaio, and Qian Yang. 2023. The Participatory Turn in AI Design: Theoretical Foundations and the Current State of Practice. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–23, Boston MA USA. ACM.

Mirko Di Lascio, Manuela Sanguinetti, Luca Anselma, Dario Mana, Alessandro Mazzei, Viviana Patti, and Rossana Simeoni. 2020. Natural Language Generation in Dialogue Systems for Customer Care. In Felice Dell'Orletta, Johanna Monti, and Fabio Tamburini, editors, *Proceedings of the Seventh Italian Conference on Computational Linguistics CLiC-it 2020*, pages 151–156. Accademia University Press.

Desmond Elliott. 2014. Towards Succinct and Relevant Image Descriptions. In *Proceedings of the Third Workshop on Vision and Language*, pages 109–111, Dublin, Ireland. Dublin City University and the Association for Computational Linguistics.

Robert E. Frederking. 2004. Grice's maxims: "do the right thing".

Abed Alhakim Freihat, Mohammed RH Qwaider, and Fausto Giunchiglia. 2018. Using Grice Maxims In Ranking Community Question Answers. In *Proceedings of the Tenth International Conference on Information, Process, and Knowledge Management, eKNOW 2018, Rome, Italy*, pages 38–43.

Terry Gaasterland, Parke Godfrey, and Jack Minker. 1992. An overview of cooperative answering. *Journal of Intelligent Information Systems*, 1(2):123–157.

Albert Gatt and Anja Belz. 2010. Introducing Shared Tasks to NLG: The TUNA Shared Task Evaluation Challenges. In Emiel Krahmer and Mariët Theune, editors, *Empirical Methods in Natural Language Generation*, volume 5790, pages 264–293. Springer Berlin Heidelberg, Berlin, Heidelberg.

Yubin Ge, Ziang Xiao, Jana Diesner, Heng Ji, Karrie Karahalios, and Hari Sundaram. 2023. What should I Ask: A Knowledge-driven Approach for Follow-up Questions Generation in Conversational Surveys. In *Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation*, pages 113–124, Hong Kong, China. Association for Computational Linguistics.

Mario Giulianelli. 2022. Towards Pragmatic Production Strategies for Natural Language Generation Tasks. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 7978–7984, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. 2017. *Towards Designing Cooperative and Social Conversational Agents for Customer Service*.

Dave Golland, Percy Liang, and Dan Klein. 2010. A Game-Theoretic Approach to Generating Spatial Descriptions. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 410–419, Cambridge, MA. Association for Computational Linguistics.

Noah D. Goodman and Michael C. Frank. 2016. Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, 20(11):818–829.

Navita Goyal, Ani Nenkova, and Hal Daumé III. 2023. Factual or Contextual? Disentangling Error Types in Entity Description Generation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8322–8340, Toronto, Canada. Association for Computational Linguistics.

Herbert Paul Grice. 1975. Logic and conversation. *Syntax and Semantics*, 3:41–58.

S. Harabagiu. 1996. Testing Gricean Constraints on a WordNet-based Coherence Evaluation System.

Ari Holtzman, Jan Buys, Maxwell Forbes, Antoine Bosselut, David Golub, and Yejin Choi. 2018. Learning to Write with Cooperative Discriminators. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1638–1649, Melbourne, Australia. Association for Computational Linguistics.

Johan F. Hoorn and Denice J. Tuinhof. 2022. A robot's sense-making of fallacies and rhetorical tropes. Creating ontologies of what humans try to say. *Cognitive Systems Research*, 72:116–130.

Laurence Robert Horn. 1972. *On the semantic properties of logical operators in English*. University of California, Los Angeles.

Dirk Hovy and Diyi Yang. 2021. The Importance of Modeling Social Factors of Language: Theory and Practice. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 588–602, Online. Association for Computational Linguistics.

Eduard Hovy. 1987. Generating natural language under pragmatic constraints. *Journal of Pragmatics*, 11(6):689–719.

Eduard H. Hovy. 1990. Pragmatics and natural language generation. *Artificial Intelligence*, 43(2):153–197.

Jennifer Hu, Sammy Floyd, Olessia Jouravlev, Evelina Fedorenko, and Edward Gibson. 2023. A fine-grained comparison of pragmatic language understanding in humans and language models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4194–4213, Toronto, Canada. Association for Computational Linguistics.

Baptiste Jacquet and Jean Baratgin. 2020. Towards a Pragmatic Model of an Artificial Conversational Partner: Opening the Blackbox. In *Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology – ISAT 2019*, pages 169–178, Cham. Springer International Publishing.

Baptiste Jacquet, Jean Baratgin, and Frank Jamet. 2018. The Gricean Maxims of Quantity and of Relation in the Turing Test. In *2018 11th International Conference on Human System Interaction (HSI)*, pages 332–338.

Baptiste Jacquet, Alexandre Hullin, Jean Baratgin, and Frank Jamet. 2019a. The Impact of the Gricean Maxims of Quality, Quantity and Manner in Chatbots. In *2019 International Conference on Information and Digital Technologies (IDT)*, pages 180–189, Zilina, Slovakia. IEEE.

Baptiste Jacquet, Olivier Masson, Frank Jamet, and Jean Baratgin. 2019b. On the Lack of Pragmatic Processing in Artificial Conversational Agents. In *Human Systems Engineering and Design*, pages 394–399, Cham. Springer International Publishing.

Prathyusha Jwalapuram. 2017. Evaluating Dialogs based on Grice's Maxims. In *Proceedings of the Student Research Workshop Associated with RANLP 2017*, pages 17–24, Varna. INCOMA Ltd.

Marten H. L. Kaas and Ibrahim Habli. 2024. Assuring AI safety: Fallible knowledge and the Gricean maxims. *AI and Ethics*.

Katarzyna Kaczmarek-Majer, Gabriella Casalino, Giovanna Castellano, Monika Dominiak, Olgierd Hryniewicz, Olga Kamińska, Gennaro Vessio, and Natalia Díaz-Rodríguez. 2022. PLENARY: Explaining black-box models in natural language through fuzzy linguistic summaries. *Information Sciences*, 614:374–399.

Assissou Kamal and Mohamed Ben Mhamed. 2023. Grice's Maxims in Moroccan EFL: A Cultural Approach through Optimality Theory. *International Journal of Linguistics, Literature and Translation*, 6(10):150–159.

Atoosa Kasirzadeh and Iason Gabriel. 2023. In Conversation with Artificial Intelligence: Aligning language Models with Human Values. *Philosophy & Technology*, 36(2):27.

Huda Khayrallah and João Sedoc. 2021. Measuring the 'I don't know' Problem through the Lens of Gricean Quantity. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5659–5670, Online. Association for Computational Linguistics.

Lea Krause, Sommerauer, Pia, and Vossen, Piek. 2022. Towards More Informative List Verbalisations. In *Joint Proceedings of the 3th International Workshop on Artificial Intelligence Technologies for Legal Documents (AI4LEGAL 2022) and the 1st International Workshop on Knowledge Graph Summarization (KGSum 2022) Co-Located with the 21st International Semantic Web Conference (ISWC 2022)*, pages 136–146, Hangzhou, China (Online). CEUR-WS.

Peter Bernard Ladkin. 2023. Involving LLMs in legal processes is risky: An invited paper. *Digital Evidence and Electronic Signature Law Review*, pages 40–46.

Raina Langevin, Ross J Lordon, Thi Avrahami, Benjamin R. Cowan, Tad Hirsch, and Gary Hsieh. 2021. Heuristic Evaluation of Conversational Agents. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15, Yokohama Japan. ACM.

Geoffrey N. Leech. 1983. *Principles of Pragmatics*. Routledge, London.

Stephen C Levinson. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.

Hengli Li, Song-Chun Zhu, and Zilong Zheng. 2023. Diplomat: A dialogue dataset for situated pragmatic reasoning. In *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*.

Ross James Lordon. 2019. *Design, Development, and Evaluation of a Patient-Centered Health Dialog System to Support Inguinal Hernia Surgery Patient Information-Seeking*. Thesis.

Rochayah Machali. 2012. Gricean maxims as an analytical tool in translation studies: Questions of adequacy. *TEFLIN*, 23(1):77–90.

Taylor Mahler, Willy Cheung, Micha Elsner, David King, Marie-Catherine de Marneffe, Cory Shain, Symon Stevens-Guille, and Michael White. 2017. Breaking NLP: Using Morphosyntax, Semantics, Pragmatics and World Knowledge to Fool Sentiment Analysis Systems. In *Proceedings of the First Workshop on Building Linguistically Generalizable NLP Systems*, pages 33–39, Copenhagen, Denmark. Association for Computational Linguistics.

Indrani Medhi Thies, Nandita Menon, Sneha Magapu, Manisha Subramony, and Jacki O'Neill. 2017. How Do You Want Your Chatbot? An Exploratory Wizard-of-Oz Study with Young, Urban Indians. In *Human-Computer Interaction - INTERACT 2017*, pages 441–459, Cham. Springer International Publishing.

Chris Mellish and Xiantang Sun. 2005. Natural Language Directed Inference in the Presentation of Ontologies. In *Proceedings of the Tenth European Workshop on Natural Language Generation (ENLG-05)*, Aberdeen, Scotland. Association for Computational Linguistics.

Erik Miehling, Manish Nagireddy, Prasanna Sattigeri, Elizabeth M. Daly, David Piorkowski, and John T. Richards. 2024. Language Models in Dialogue: Conversational Maxims for Human-AI Interactions. *Preprint*, arxiv:2403.15115.

Yunju Nam, Hyenyeong Chung, and Upyong Hong. 2023. Language Artificial Intelligences' Communicative Performance Quantified Through the Gricean Conversation Theory. *Cyberpsychology, Behavior, and Social Networking*, 26(12):919–923.

Eric W. T. Ngai, Maggie C. M. Lee, Mei Luo, Patrick S. L. Chan, and Tenglu Liang. 2021. An intelligent knowledge-based chatbot for customer service. *Electronic Commerce Research and Applications*, 50:101098.

Antinus Nijholt. 2011. No Grice: Computers that Lie, Deceive and Conceal. In *Proceedings 12th International Symposium on Social Communication*, pages 889–895. Centre for Applied Linguistics.

Manuel Ocaña, David Chapela-Campa, Pedro Álvarez, Noelia Hernández, Manuel Mucientes, Javier Fabra, Ángel Llamazares, Manuel Lama, Pedro A. Revenga, Alberto Bugarín, Miguel A. García-Garrido, and Jose M. Alonso. 2022. Automatic linguistic reporting of customer activity patterns in open malls. *Multimedia Tools and Applications*, 81(3):3369–3395.

Kaseem Oladimeji Olaniyi and Josephine Olushola Oyinbo. 2021. Gricean Pragmatics and the English Language in Nigeria. *Journal of Second and Multiple Language Acquisition-JSMULA*, pages 226–240.

Silviu Oprea, Steven Wilson, and Walid Magdy. 2021. Chandler: An Explainable Sarcastic Response Generator. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 339–349, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Amit Kumar Pandey, Rodolphe Gelin, and AMPSH Robot. 2018. Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3):40–48.

Laura Panfili, Steve Duman, Andrew Nave, Katherine Phelps Ridgeway, Nathan Eversole, and Ruhi Sarikaya. 2021. Human-AI interactions through a Gricean lens. *Proceedings of the Linguistic Society of America*, 6(1):288–302.

Dojun Park, Jiwoo Lee, Hyeyun Jeong, Seohyun Park, and Sungeun Lee. 2024. Pragmatic Competence Evaluation of Large Language Models for Korean. *Preprint*, arxiv:2403.12675.

M. Pereira-Fariña, Luka Eciolaza, and Gracian Trivino. 2012. Quality Assessment of Linguistic Description of Data. In *Proceeding of the 16th Conference on Fauzzy Logic and Technologies*, pages 608–613.

Chiara Barattieri di San Pietro, Federico Frau, Veronica Mangiaterra, Valentina Bambini, and Chiara Barattieri di San Pietro. 2023. The pragmatic profile of ChatGPT: Assessing the communicative skills of a conversational agent. *Sistemi intelligenti*, (2/2023).

Albert Pritzkau, Julia Waldmüller, Olivier Blanc, Michaela Geierhos, and Ulrich Schade. 2023. Current language models' poor performance on pragmatic aspects of natural language. In *Working Notes of FIRE 2023 - Forum for Information Retrieval Evaluation (FIRE-WN 2023), Goa, India, December 15-18, 2023*, volume 3681 of *CEUR Workshop Proceedings*, pages 159–169. CEUR-WS.org.

Mohammed R. H. Qwaider, Abed Alhakim Freihat, and Fausto Giunchiglia. 2017. TrentoTeam at SemEval-2017 Task 3: An application of Grice Maxims in Ranking Community Question Answers. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 271–274, Vancouver, Canada. Association for Computational Linguistics.

Roni Rabin, Alexandre Djerbetian, Roee Engelberg, Lidan Hackmon, Gal Elidan, Reut Tsarfaty, and Amir Globerson. 2023. Covering Uncommon Ground: Gap-Focused Question Generation for Answer Assessment. *Preprint*, arxiv:2307.03319.

Ehud Reiter. 1990. The Computational Complexity of Avoiding Conversational Implicatures. In *28th Annual Meeting of the Association for Computational Linguistics*, pages 97–104, Pittsburgh, Pennsylvania, USA. Association for Computational Linguistics.

Douglas Robinson. 2002. *Performative Linguistics: Speaking and Translating as Doing Things with Words*. Routledge, London.

Philipp Sadler and David Schlangen. 2023. Pento-DIARef: A Diagnostic Dataset for Learning the Incremental Algorithm for Referring Expression Generation from Examples. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 2106–2122, Dubrovnik, Croatia. Association for Computational Linguistics.

Mohammad Saleh Sanatifar and Mohammad Sadegh Kenevisi. 2017. The idea of faithfulness and reformulation of the Gricean maxims for the needs of translation. *FORUM. Revue internationale d'interprétation*

et de traduction / International Journal of Interpretation and Translation, 15(1):67–84.

Manuela Sanguinetti, Alessandro Mazzei, Viviana Patti, Marco Scalerandi, Dario Mana, and Rossana Simeoni. 2020. Annotating Errors and Emotions in Human-Chatbot Interactions in Italian. In *Proceedings of the 14th Linguistic Annotation Workshop*, pages 148–159, Barcelona, Spain. Association for Computational Linguistics.

Ayse Pinar Saygin and Ilyas Cicekli. 2002. Pragmatics in human-computer conversations. *Journal of Pragmatics*, 34(3):227–258.

Matthias Scheutz, Ravenna Thielstrom, and Mitchell Abrams. 2022. Transparency through Explanations and Justifications in Human-Robot Task-Based Communications. *International Journal of Human–Computer Interaction*, 38(18-20):1739–1752.

S. M. Seals and Valerie L. Shalin. 2023. Discourse over Discourse: The Need for an Expanded Pragmatic Focus in Conversational AI. *Preprint*, arxiv:2304.14543.

Vidya Setlur and Melanie Tory. 2022. How do you Converse with an Analytical Chatbot? Revisiting Gricean Maxims for Designing Analytical Conversational Behavior. In *CHI Conference on Human Factors in Computing Systems*, pages 1–17, New Orleans LA USA. ACM.

Avinash Kumar Singh, Neha Baranwal, Kai-Florian Richter, Thomas Hellström, and Suna Bensch. 2021. Verbal explanations by collaborating robot teams. *Paladyn, Journal of Behavioral Robotics*, 12(1):47–57.

Jonas Sjöbergh and Kenji Araki. 2008. What is poorly Said is a Little Funny. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco. European Language Resources Association (ELRA).

Mohammad Sorower, Janardhan Doppa, Walker Orr, Prasad Tadepalli, Thomas Dietterich, and Xiaoli Fern. 2011. Inverting Grice' s Maxims to Learn Rules from Natural Language Extractions. In *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc.

Settaluri Lakshmi Sravanthi, Meet Doshi, Tankala Pavan Kalyan, Rudra Murthy, Pushpak Bhattacharyya, and Raj Dabre. 2024. PUB: A Pragmatics Understanding Benchmark for Assessing LLMs' Pragmatics Capabilities. *Preprint*, arxiv:2401.07078.

Somayajulu G. Sripada, Ehud Reiter, Jim Hunter, and Jin Yu. 2003. Generating English summaries of time series data using the Gricean maxims. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '03, pages 187–196, New York, NY, USA. Association for Computing Machinery.

Yufei Tao, Ameeta Agrawal, Judit Dombi, Tetyana Sydorenko, and Jung In Lee. 2024. ChatGPT Roleplay Dataset: Analysis of User Motives and Model Naturalness. *Preprint*, arxiv:2403.18121.

Marta Tatu and Dan Moldovan. 2012. A Tool for Extracting Conversational Implicatures. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 2708–2715, Istanbul, Turkey. European Language Resources Association (ELRA).

Maitreyee Tewari, Suna Bensch, Thomas Hellström, and Kai-Florian Richter. 2020. Modelling Grice's Maxim of Quantity as Informativeness for Short Text. In *ICLLL 2020 : The 10th International Conference in Languages, Literature, and Linguistics, Japan, November 6-8, 2020*, pages 1–7.

Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, YaGuang Li, Hongrae Lee, Huaixiu Steven Zheng, Amin Ghafouri, Marcelo Menegali, Yanping Huang, Maxim Krikun, Dmitry Lepikhin, James Qin, Dehao Chen, Yuanzhong Xu, Zhifeng Chen, Adam Roberts, Maarten Bosma, Vincent Zhao, Yanqi Zhou, Chung-Ching Chang, Igor Krivokon, Will Rusch, Marc Pickett, Pranesh Srinivasan, Laichee Man, Kathleen Meier-Hellstern, Meredith Ringel Morris, Tulsee Doshi, Renelito Delos Santos, Toju Duke, Johnny Soraker, Ben Zevenbergen, Vinodkumar Prabhakaran, Mark Diaz, Ben Hutchinson, Kristen Olson, Alejandra Molina, Erin Hoffman-John, Josh Lee, Lora Aroyo, Ravi Rajakumar, Alena Butryna, Matthew Lamm, Viktoriya Kuzmina, Joe Fenton, Aaron Cohen, Rachel Bernstein, Ray Kurzweil, Blaise Aguera-Arcas, Claire Cui, Marian Croak, Ed Chi, and Quoc Le. 2022. LaMDA: Language Models for Dialog Applications. *Preprint*, arxiv:2201.08239.

Kees van Deemter. 2002. Generating Referring Expressions: Boolean Extensions of the Incremental Algorithm. *Computational Linguistics*, 28(1):37–52.

Chris van der Lee, Albert Gatt, Emiel van Miltenburg, and Emiel Krahmer. 2021. Human evaluation of automatically generated text: Current trends and best practice guidelines. *Computer Speech & Language*, 67:101151.

Sebastian Varges, Heike Bieler, Manfred Stede, Lukas C. Faulstich, Kristin Irsig, and Malik Atalla. 2012. SemScribe: Natural Language Generation for Medical Reports. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 2674–2681, Istanbul, Turkey. European Language Resources Association (ELRA).

Jette Viethen and Robert Dale. 2006. Algorithms for Generating Referring Expressions: Do They Do What People Do? In *Proceedings of the Fourth International Natural Language Generation Conference*, pages 63–70, Sydney, Australia. Association for Computational Linguistics.

Adam Vogel, Max Bodoia, Christopher Potts, and Daniel Jurafsky. 2013. Emergence of Gricean Maxims from Multi-Agent Decision Theory. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1072–1081, Atlanta, Georgia. Association for Computational Linguistics.

Wolfgang Wahlster. 2023. Understanding computational dialogue understanding. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 381(2251):20220049.

Ludwig Wittgenstein. 1953. Philosophical investigations. *Basil & Blackwell, Oxford*.

Matthias Wölfel, Mehrnoush Barani Shirzad, Andreas Reich, and Katharina Anderer. 2024. Knowledge-Based and Generative-AI-Driven Pedagogical Conversational Agents: A Comparative Study of Grice's Cooperative Principles and Trust. *Big Data and Cognitive Computing*, 8(1):2.

Ziang Xiao, Michelle X. Zhou, Q. Vera Liao, Gloria Mark, Changyan Chi, Wenxi Chen, and Huahai Yang. 2020. Tell Me About Yourself: Using an AI-Powered Chatbot to Conduct Conversational Surveys with Open-ended Questions. *ACM Transactions on Computer-Human Interaction*, 27(3):1–37.

R. Michael Young. 1999. Using Grice's maxim of Quantity to select the content of plan descriptions. *Artificial Intelligence*, 115(2):215–256.

Shisen Yue, Siyuan Song, Xinyuan Cheng, and Hai Hu. 2024. Do Large Language Models Understand Conversational Implicature – A case study with a chinese sitcom. *Preprint*, arxiv:2404.19509.

Zilong Zheng, Shuwen Qiu, Lifeng Fan, Yixin Zhu, and Song-Chun Zhu. 2021. GRICE: A Grammar-based Dataset for Recovering Implicature and Conversational rEasoning. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 2074–2085, Online. Association for Computational Linguistics.

A Appendix

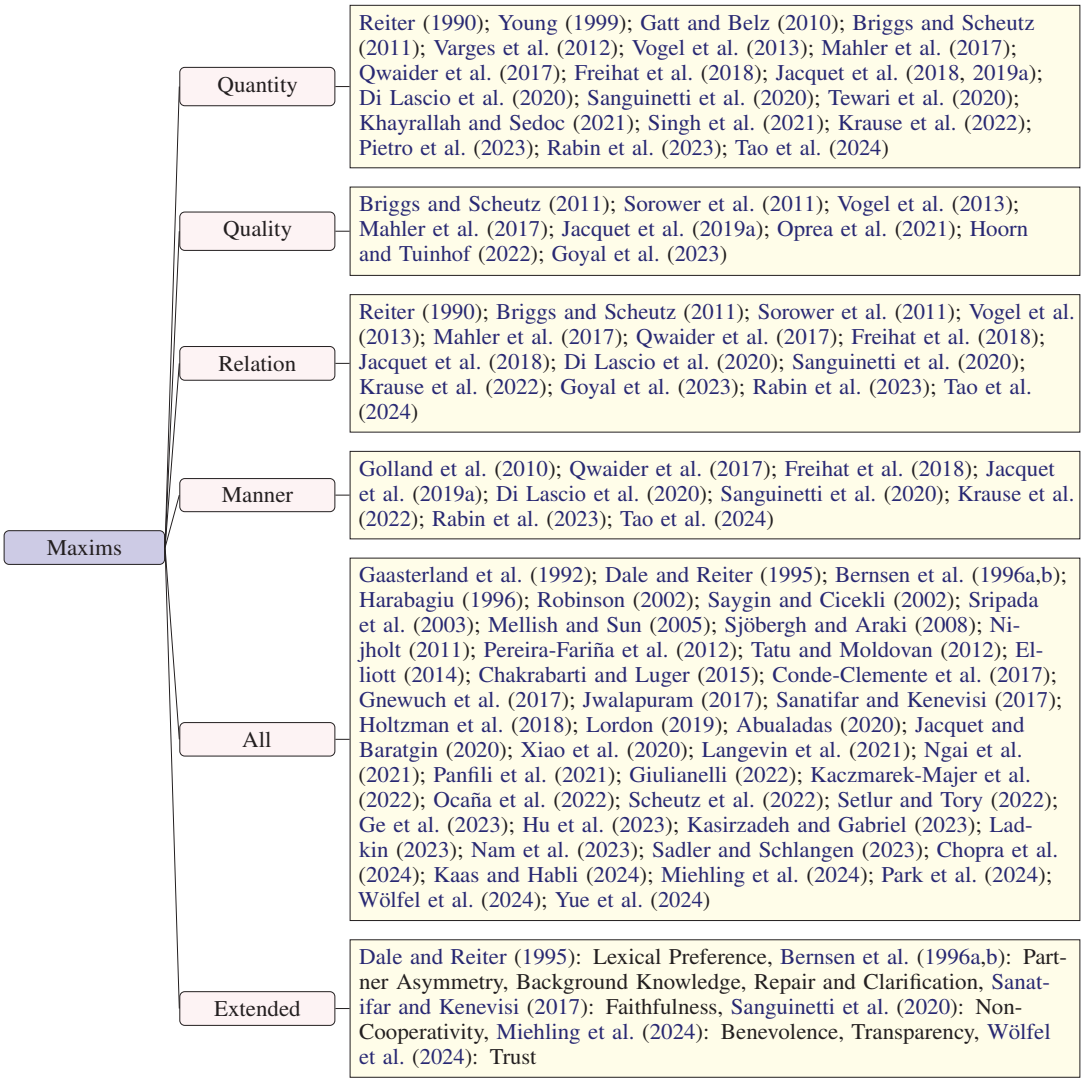

Figure 2: The figure categorises papers based on the specific Gricean maxims they address. Some papers mention all or other maxims but focus only on a subset for in-depth analysis or application. In cases where papers elaborate on additional maxims beyond the standard ones, these are included in *Extended* unless the focus is part of a larger evaluation.