ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

JOURNAL OF APPLIED LINGUISTICS AND TESOL

Vol.8. No.3.2025

INTEGRATING PRAGMATICS WITH LARGE LANGUAGE MODELS: A MULTILINGUAL APPROACH TO CONTEXTUAL UNDERSTANDING

Luqman Manzoor
MPhil applied linguistics, Government College University Faisalabad
Email: luqmanmanzoor875@gmail.com
Rubab Gull
MPhil Applied Linguistics, Government College University Faisalabad
Email: gulrubab73@gmail.com
Zakra Nadeem
M.phil English Literature, Riphah International University Faisalabad
Email: zakranadeem67@gmail.com

DOI: https: doi.org/10.5281/zenodo.17221857

Abstract
The state of the art performance of the large language models (LLM) has transformed the natural language processing system by delivering essential results in various linguistic activities such as translation, summarization, and conversational interactions. However, despite their fine acting, such models possess a single significant weakness the lack of pragmatic depth. God-given human communication depends on pragmatics or the meaning and context within the environment, mission, and cultural regulations that define this meaning. It does not imply that the indirect speech acts, implicatures, politeness strategies, and culturally-aware phrases cannot be misunderstood by the LLMs, which are highly powerful in syntax and semantics. This has been blown out of proportion where in the multilingual setup, one and the same utterances could be pragmatically significant in one linguistic community and even in another linguistic community.
A mixture of practical principles in the LLM designs is the solution of this challenge. The paper will propose a multilingual practice, which is founded on the contextual conceptualization of languages, on the pragmatics and intercultural communication theories. The research is scheduled to expand the functions of the LLCs with pragmatic consciousness by experimenting with pragmatically scribbled corpora, as well as creating an inventory of testing measures which quantify the aptitude of the interpreters to be circumstantially conscious. Further, the paper considers the strategies of fine-tuning and comparison of human judgment and the output of LLM to evaluate pragmatic fidelity.
Hopefully, one may realize that the contributions are triple to demonstrate necessity to introduce pragmatic reasoning into the LLMs, to underline the necessity of multilingualism in the context of forming the context awareness, andto propose the paradigms of pragmatic incorporation that could enhance cross-cultural communication. Long-term, the proposed study is hoping that the result will be an even contextually aware, culturally sensitive, and more trustworthy LLM that is closer to human-like language competence.

Keywords
Pragmatics; Large Language Models; Multilingualism; Contextual Understanding; Cross-Cultural Communication; Natural Language Processing; Computational Linguistics

1. Introduction

Over the last ten years, Large Language Models (LLMs) have become a research and application platform of natural language processing (NLP). Models such as GPT-4 (OpenAI, 2023), PaLM

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

(Chowdhery et al., 2022), and LLaMA (Touvron et al., 2023) have demonstrated impressive text generation, translation, summarization, and dialogue systems performance based on huge datasets and billions of parameters. These models operate by providing the estimates on the most likely word order using a certain prompt, and the findings are consistent and far more natural. This is evidenced by their quick spread in academia, industry and the society and indicates that they can revolutionize human-computer interaction in a new manner.

Even with these improvements, it is important to note that the LLMs are still facing major problems when it comes to representing the finer aspects of human communication. Systematic and semantics are sufficiently addressed, but pragmatics, which involve the way meaning is built in the context, intent, and culture, is still a lingering weakness. Classic pragmatics theories, including Cooperative Principle introduced by Grice (1975), Speech Act Theory by Austin (1962) and Relevance Theory introduced by Sperber and Wilson (1986) focus on the extent to which communication relies on inference beyond literal word meanings. As an example, a speaker can express himself using the word it is cold in here which may have a connotation meaning that the speaker wants the window to be closed instead of a statement of fact. The weak points of LLMs are such indirectness, implicatures, politeness strategies, irony, and culturally encoded messages (Haugh and Chang, 2019; Zhou et al., 2023).

The problem is intensified in multilingual situations. Pragmatic norms are quite different among different linguistic communities and, therefore, context-sensitive interpretation is the key to successful communication. Studies on cultural pragmatics show that the same utterances may have different meanings in different languages (Blum-Kulka, 1989; House, 2009). As an example, the way of being polite in English will usually be indirect, but in Russia it is more direct and in Japan it is more liberal with honorific and subordinate relationships (Ide, 1989; Wierzbicka, 2003). Unmindful of such practical diversity, the results that LLMs gives may be linguistically correct and pragmatically incorrect and may cause miscommunication or cultural insensitivity.

Current literature has started covering certain issues of pragmatics in AI, yet there are still a lot of gaps. Although there are improvements in the contextual embeddings and fine-tuning methods to enhance coherence and discourse awareness (Devlin et al., 2019; Liu et al., 2019), these approaches fail to consider pragmatic reasoning. Certain efforts to provide the LLAM with external body of knowledge or dialogue history have been encouraging (Zhou et al., 2023), but the systematic consideration of pragmatic principles, especially in a multilingual context, is poorly studied. In addition, the assessment of LLMs, as a rule, does not revolve around pragmatic competence, giving preference to syntactic accuracy or semantic accuracy. This is an absence of focus on pragmatics that presents a significant gap in theory and practice to restrain the functioning of AI systems within a culturally diverse communication environment.

This study fills that gap by suggesting a framework of considering pragmatics in LLM using a multilingual perspective. Based on pragmatic theory, intercultural communication studies, and recent developments in computational linguistics, the study is aimed at investigating how LLMs can go beyond literal interpretation and towards context-sensitive, culturally sensitive interpretation. This study is formulated with a triple objective, namely, (1) to understand how pragmatic principles can be made operative within the context of LLM, (2) to examine the contribution of multilingualism in the formation of pragmatic interpretation, and (3) to advance

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

and experiment on frames of assessing and improving pragmatic competence in the context of LLM.

The research questions that will be used to conduct this study are:

1. How far can existing LLMs understand pragmatic phenomena like implicature, politeness strategies and indirect speech acts in different languages?
2. What are some of the differences of pragmatic norms between linguistic and cultural contexts and what are the challenges to multilingual LLM?
3. Which computational strategies may be developed to introduce pragmatic reasoning into LLMs and evaluate their work in a multilingual context?

This way, the research can not only advance the field of NLP, but the overall objective of the research is to promote cross-cultural communication in the era of AI. Pragmatic enrichment, particularly in multi-lingual contexts is an important developmental process towards the development of models that are technically competent as well as socially and culturally sensitive.

2. Literature Review

2.1 Theoretical pragmatics

As a branch of linguistics, pragmatics looks into meaning construction outside of literal semantics and deals with the interaction between language, context and speaker intention. The Speech Act Theory by Austin (1962) had the conceptualization that utterances are agents, and include requesting, apologising, or commanding things instead of describing things. Based on this, Searle (1969) has explained the categories of speech acts and has stressed on the significance of the intention of speakers in the creation of meaning.

Cooperative Principle and conversational maxims (quantity, quality, relation, and manner) presented by Grice (1975) offered an insight into comprehending implicature, in which listeners deduce unspoken meanings through mutual assumptions of cooperation. Indicatively, a question like, What was it like at the conference, an answer like, the food was good, will be taken to mean that the performance of the conference was poor. The contextual inference required in such practical arguments is beyond what is literally meant in the words.

Relevance Theory (Sperber and Wilson, 1986/1995) took the pragmatic theory further by claiming that the process of communication is directed by the optimal relevance search which needed the optimal compromises between cognitive processing and the effect of context. This school of thought emphasises the dynamism of interpretation of utterances by the listeners depending on previous knowledge, the history of the discourse and the culture.

These theories are the pillars to support the reason why pragmatics cannot be ignored in human communication. However, they also indicate the limitations of computational models: although lexical and syntactic patterns can provide the semantic meaning, pragmatic interpretation requires contextual inferences, cultural knowledge, and knowledge of the intention of a speaker (Levinson, 1983; Thomas, 1995).

2.2 LLMs and Contextual Challenges

LLMs like GPT-3 (Brown et al., 2020), PaLM (Chowdhery et al., 2022), and LLaMa (Touvron et al., 2023) are extraordinary in terms of their capacity to produce coherent text in a variety of tasks (question answering, creative writing, and others). These models take advantage of transformer architectures (Vaswani et al., 2017), which rely on attention mechanisms to extract long-range text dependencies. Prior acquaintance with large-scale corpora provides them with a large amount of statistical experience with linguistic patterns.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

Nevertheless, LLMs are good at semantic capture and generation of fluent sentences but fail to capture pragmatic signals. Research has revealed that these models are not good at indirect requests, sarcasm, irony, and context-specific politeness (Haugh and Chang, 2019; Huang et al., 2023). An example is when someone is presented with ironic messages, like telling him or her that he/she did a great job after a mistake, very often the model interprets the message as being completely literal.

The other weakness is that the grounded world knowledge of LLMs is not applicable in interactional environments. In spite of certain advances achieved with the help of retrieval mechanisms (Lewis et al., 2020) and dialogue fine-tuning (Adiwardana et al., 2020), pragmatic interpretation is not possible only with the availability of factual information but knowledge of the social norms, speaker intentions, and cultural frames.

Also, present-day evaluation scales, including BLEU (Papineni et al., 2002) or ROUGE (Lin, 2004), focus on lexical overlap and semantic similarity and not pragmatic fidelity. Therefore, models can produce syntactically sound translations or summaries, but they can still fail to address the pragmatic intent of the speaker (Ji et al., 2023). This inconsistency brings out the necessity of practical incorporation in model frameworks and assessment frameworks.

2.3 Multilingual Pragmatics

Cultural and linguistic norms necessarily influence pragmatics and therefore multilingual perspectives should be considered in AI research. Cross-cultural pragmatics researches have shown that speech acts, politeness strategies, and implicatures differ greatly across languages (Blum-Kulka, 1989; House, 2009). As the case in point, indirect forms are commonly borrowed by English speakers to make their requests less offensive, but at the same time Russian speakers can also employ more direct forms and still be viewed as not rude (Wierzbicka, 2003). In the same way, Japanese communication is characterized by the use of honorifics and hierarchical sensitivity (Ide, 1989) and Arabic is more characterized by formulaic expressions of politeness that are related to religious or cultural values (Al-Khatib, 2001).

Multilingual LLM have special challenges in such diversity. The translation of the pragmatically rich utterances is the case when a literal translation leads to a translation product that is grammatically correct and pragmatically unsuitable. As an example, directing English indirect demands to Russian in the same way indirectly, there is a risk of indirectness coming out as excessively formal and sarcastic. This effect is referred to as pragmatic transfer and it highlights the challenge of maintaining the pragmatic parallel in different languages (Kecskes, 2014).

Intercultural pragmatics also indicate in research that pragmatic competence is not solely linguistic, but also social-cultural (Ishihara and Cohen, 2010; Taguchi, 2019). The speakers use common cultural knowledge to derive implicatures, humor and politeness. In the absence of the modeling of such knowledge, multilingual LLMs can be reinforced in stereotypes, misunderstanding intent, or produce culturally blind responses.

2.4 Efforts to Incorporate Pragmatics into AI

There is a recent push to investigate pragmatic integration within AI systems. Among them, there is the method of annotating corpora with pragmatic characteristics, i.e. speech act and implicature type, or politeness (Bunt et al., 2010; Danescu-Niculescu-Mizil et al., 2013). Recognizing pragmatic phenomena with training on such enriched datasets can be improved, but at large-scale levels of pragmatic annotation, it is still resource-intensive.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

The other approach is that of basing models upon discursive and social engagement. Conversation systems like Meena (Adiwardana et al., 2020) and BlenderBot (Roller et al., 2021) strive to encode the long-term context of the conversation, but still, they fail to capture less pronounced pragmatic information. Other researchers test reinforcement learning based on human control (Christiano et al., 2017; Ouyang et al., 2022), which improves pragmatic performance indirectly by matching outputs with the expectations of users.

Multimodal integration also has more opportunities. Models would be able to estimate pragmatic and reasoning by using nonverbal cues, including prosody, facial expression, and gestures (Li et al., 2021). Nevertheless, the field remains immature and experiences the problem of scalability and cultural diversity.

Regardless of these developments, there are still no elaborate frameworks of pragmatic integration. According to Zhou et al. (2023), pragmatic reasoning needs both annotated data and new architectures that can make inferences outside of text. Likewise, as pointed out by Haugh (2018), pragmatic meaning is a result of interactional processes, which can hardly be recreated by the static pre-trained models. The above observations highlight the importance of systematic strategies that would liaise between linguistic theory, cross-cultural pragmatics, and computational design.

2.5 Research Gap

The review highlights several key gaps in current research:

1. **Limited pragmatic competence in LLMs**: While models excel at syntax and semantics, they frequently misinterpret implicatures, politeness strategies, and indirectness (Huang et al., 2023).
2. **Insufficient focus on multilingual pragmatics**: Most studies evaluate LLM performance in English or high-resource languages, neglecting how pragmatic norms differ across linguistic and cultural contexts (House, 2009; Wierzbicka, 2003).
3. **Lack of pragmatic evaluation metrics**: Current benchmarks focus on lexical or semantic accuracy, leaving pragmatic interpretation underexamined (Ji et al., 2023).
4. **Fragmented integration efforts**: Existing attempts to embed pragmatics into AI remain isolated, lacking a unifying framework that systematically addresses multilingual and intercultural dimensions (Zhou et al., 2023).

Thus, there is a pressing need for research that not only identifies these shortcomings but also develops computational approaches to integrate pragmatics into LLMs. A multilingual lens is particularly important, as it ensures that models can function effectively across diverse cultural and linguistic settings, thereby promoting more authentic and contextually appropriate communication.

2.6 Research Problem & Objectives

No matter how impressive the results that Large Language Models (LLMs) have managed to deliver in the field of natural language processing are, their failure in pragmatic competence can still be noticed. The current systems are highly competent at creating syntactically accurate and semantically sound text, yet do not comprehend deeper layers of meaning that depend on context, the intent of the speaker and cultural standards (Huang et al., 2023; Haugh, 2018). Indirect request, conversational implicatures, sarcasm are some of the examples which are misunderstood and deliver the result that appears to be outwardly correct, but pragmatically inappropriate. This is a disadvantage since it reduces the communicative abilities of the LLMs to

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

that of a human being and reduces their applicability in real-life activities such as translation, intercultural conversation and conversational AI.

The issue gains even greater importance in the multilingual setting where pragmatic norms differ considerably among cultures. Words used to mean something polite in particular language might be perceived as impolite or too formal when translated directly into different language (Wierzbicka, 2003; House, 2009). As they are insensitive to such cross-cultural differences, multilingual LLMs are likely to distort the intent of the speakers or produce culturally inappropriate answers. This is a matter of great concern especially in situations where AI-mediated communication is characterized by a wide range of users, including international business, international relations, or international education (Kecskes, 2014).

Despite certain improvements presented by pragmatic annotation (Bunt et al., 2010) and reinforcement learning based on human feedback (Ouyang et al., 2022), a systematic approach to conceptualize pragmatics in the design and evaluation of LLM does not exist so far. Existing metrics focus on semantic fidelity and avoid discussing an important dimension of meaning, which is the prevailing aspect (Ji et al., 2023). This then means that there exists an outstanding gap in research: the necessity to operationalize pragmatic theories in computational models in a manner that would explain multilingual and intercultural diversity.

2.7 Research Objectives

The aims of this research are to fill the mentioned gap because the research intends to accomplish:

1. To instantiate practical principles (e.g., implicature, politeness, indirectness) in the LLM architectures.
2. To explore how multilingualism influences the development of pragmatic interpretation and what difficulties it presents to cross-cultural communications.
3. To model and test computational models incorporating pragmatic reasoning in LLMs, with pragmatically annotated multilingual corpora.
4. To suggest new assessment scales to evaluate pragmatic fidelity as well as semantic accuracy.

Through such accomplishments, the research project will contribute to the evolution of linguistically proficient and pragmatically competent and culturally sensitive LLMs.

3. Methodology

The paper shall be a mixed-method work in order to investigate how pragmatics could be integrated into Large Language Models (LLM) in a multilingual view. It is a mixture of corpus based analysis, experimental fine-tuning of models, and the construction of pragmatic evaluation measures. The reason why such design is required is that theoretical knowledge of pragmatics and practical issues involving computational modeling are considered.

3.1 Research Design

The study is an exploratory sequential study. It starts with a qualitative analysis of pragmatically complicated multilingual evidence in an attempt to describe the patterns to misinterpretation of the current outputs of LLM. The second step makes use of the findings to refine the experimental application of pragmatically annotated corpus based on LLMs. The final thing would be to measure the outputs of the models quantitatively based on innovative pragmatic measures and compare to human judgments. This is a multi-layered approach that can be used to obtain an overview of the issue (Creswell and Plano Clark, 2018).

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

3.2 Data Collection

3.2.1 Pragmatically Annotated Corpora

The research will apply the already existing pragmatically annotated datasets, including the ISO 24617-2 Dialogue Act Annotation Scheme (Bunt et al., 2010), or corpora marked-up with politeness and implicature signs (Danescu-Niculescu-Mizil et al., 2013). Further information will be obtained by means of multilingual parallel corpora (e.g., Europarl, OPUS) to study cross-linguistic variation. In the areas where it lacks, small-scale annotation work will be done, with the languages of case study, English, Russian, Japanese, and Arabic, because they represent different pragmatic traditions (House, 2009; Ide, 1989; Wierzbicka, 2003).

3.2.2 Human Judgment Data

To supplement corpus analysis, model outputs will be evaluated by human assessors who will possess multilingual ability with respect to pragmatic appropriateness. This makes the evaluations to be based on real cultural knowledge and not on the surface-level linguistic accuracy.

3.2.3 Experimental Procedure

The first test of the LLaMA-2 will be on pragmatically challenging data, including the ability to comprehend an indirect request, sarcasm, or culturally sensitive politeness strategies. Their results will be checked with those of human interpretation in order to create a baseline performance.

3.2.4 Refinements with Practicable Data.

Trained pragmatically annotated corpora will then be used to fine-tune selected LLMs. The supervised learning on the labelled examples of speech acts, implicatures, and strategies of politeness will be part of the fine-tuning process. Cross-linguistic variation will be highlighted to make sure that the models do not extrapolate the pragmatic norms between languages.

3.2.5 Evaluation

The assessment phase will use both conventional measures and new pragmatic based measures. Even though BLEU (Papineni et al., 2002) and ROUGE (Lin, 2004) will evaluate semantic overlap, other measures will be used to evaluate pragmatic fidelity, including:

1. Implicature Recognition Rate (IRR): model capability to get intended meaning out of literal meaning.
2. Politeness Preservation Score (PPS): skill to retain multilingual translations or multilingual conversations that employ culturally appropriate politeness strategies.
3. Indirect Speech Act Accuracy (ISAA): the ability to understand indirect requests or commands.

The reliability of these metric will be proven by the comparative analysis with human judgments.

3.3 Data Analysis

Cases of success in models that do not produce pragmatic meaning and those that are successful will be analyzed qualitatively. As an example, the questions of whether irony is appropriately interpreted or whether the strategies of politeness are maintained in the process of translation can be investigated. The quantitative analysis will also entail computing precision, recall, and F1 scores of pragmatic phenomena and the statistical comparisons (e.g., t-tests, ANOVA) between the baseline and fine-tuned models.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

3.3.1 Ethical Considerations

As the research is the interaction with multilingual and multicultural information, ethical principles become the key. Biases in corpora training can support stereotypes or distort cultural conventions (Bender et al., 2021). The human evaluators will be chosen with great care in regard to cultural competence and their annotations will be anonymized in order to safeguard privacy. Besides, the study will make sure that the outputs are not just pragmatically suitable but also respectful of the diversity of cultures in order not to strengthen the discriminatory or harmful discourse.

3.3.2 Limitations

A number of methodological shortcomings have to be mentioned. To begin with, pragmatic annotation is not a light resource, and annotated corpora of large scale are resources of a rare kind. This limits the fine-tuning ability and can cause a sampling bias. Second, pragmatic competence implies interaction, whereas LLMs usually rely on fixed textual inputs; this could also restrict their capability of replicating human-like inference to the full extent (Haugh, 2018). Third, multilingual coverage will also be selective, i.e. not based on an international collection but on a number of case study languages. Still, the mixed-methods approach overcomes these shortcomings by involving data corpus, human assessment and experimental modeling.

4. Discussion / Analysis

The overall results of the literature and the methodological framework suggest the necessity of implementing pragmatics in Large Language Models (LLMs). This section describes my analysis of the challenges, possible solutions, and general implications of pragmatic enrichment to multilingual contextual understanding. It is based on four main themes, such as pragmatic enrichment as a requirement of LLMs, cross-cultural challenges to multilingual pragmatics, ethical and societal issues, and integration frameworks.

4.1 Pragmatic Enrichment as a Requirement to LLMs

The performance of LLMs is significantly worse in capturing pragmatic nuances: they have always been found to fail in situations where meaning is through indirectness, implicatures, or cultural conventions (Huang et al., 2023). It is no surprise that this is limited: semantic meaning may be learnt under statistical co-occurrences in training data, whereas pragmatic meaning needs to be inferred using contextual information and common cultural knowledge (Levinson, 1983; Thomas, 1995). To use the example of Can you pass the salt, this is a request and not an inquiry as to whether a person can pass the salt or not. However, without pragmatic explicit grounding, YLLMs may consider it to be the latter.

The above methodology proves the idea that pragmatic annotation and human judgment can reveal these vices. Pragmatically complex tasks can be tested on the baseline models to measure to what extent they are based on the surface-level semantics instead of the contextual reasoning. Even in the initial research, it is already known that LLMs are prone to misunderstanding sarcasm and politeness techniques (Zhou et al., 2023). Such failures may undermine the confidence in AI systems, especially in such sensitive fields as healthcare, education, and intercultural mediation.

Hence, pragmatic enrichment is not an optional refinement, but a requirement of obtaining human-like communication competence. In its absence, LLMs will run the danger of generating socially tone-deaf but fluency-in-the-real-world outputs, which undermines their utility in practice.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

4.2 Multilingual Pragmatics Cross-Cultural Complexities

One of the key conclusions made by intercultural pragmatics studies is that pragmatic norms vary in linguistic communities (Blum-Kulka, 1989; House, 2009). These are challenges peculiar to multilingual LLM. An example is that when English prefers indirect methods to express politeness, Russian prefers directness, Japanese focuses on honorific and Arabic prefers to incorporate politeness in religious phrases (Ide, 1989; Wierzbicka, 2003; Al-Khatib, 2001).

By trying to translate or interpret between languages without taking these differences into consideration, LLMs are likely to commit pragmatic transfer errors (Kecskes, 2014). As an example, a direct translation of the English words Could you possibly open the window into Russian can sound too formal, or even sarcastic. On the same note, the failure to produce correct honorifics in Japanese would lead to rudeness.

It is especially useful in the multilingual design of the methodology that refers to annotated corpora, in the case of English, Russian, Japanese, and Arabic. It provides the opportunity to compare the results of cross-cultural pragmatic phenomenon by the performance of LLMs. Owing to fine-tuning of the pragmatically annotated multilingual data, it becomes possible to start distinguishing between language-specific and more universal pragmatic strategy.

However, challenges remain. Pragmatic rules do not remain the same, instead, they change with each generation, digital communication, and cultural hybridities (Taguchi, 2019). The dynamism is only possible to capture by constantly updating corpora and evaluation frameworks. Additionally, the low-resource languages are another challenge: whereas English and Japanese might feature a comparatively rich set of pragmatic studies, a lot of languages do not have annotated corpora at all, and multilingual coverage remains incomprehensive (Bender et al., 2021). It is highly important to curb these discrepancies so that pragmatic integration would not add to linguistic inequalities.

4.3 Ethical and Social Concerns

Pragmatics is not only technically challenging to integrate into the LLMs, but it is also ethically required. Any misunderstanding of pragmatic meaning may create misunderstanding, insult or even damage. As an example, a politeness strategy translated mistranslatively in the case of diplomatic communication might be perceived unintentionally as disrespect. Healthcare providers may fail to provide care to a patient because of the indirect manifestation of pain or discomfort (Haugh, 2018).

Another burning issue is prejudice. The pragmatic norms are usually mirrors of hierarchies and gender roles, and authority (Mills, 2003). When the LLM is exposed to these norms without question, they would be able to reproduce or even enhance discriminatory practices. As an illustration, models may imitate gendered patterns of politeness that debase women in some linguistic situations. In order to counter these risks, the methodology includes human evaluators that are culturally competent such that pragmatic decisions are based on genuine and diverse viewpoints.

4.4 Models of Pragmatics Integration into LLMs

Based on the literature and the findings related to the methods, various structures of incorporating pragmatics into the LLMs arise.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

a) Pragmatically Annotated Training

The initial step will be to train pragmatically enhanced datasets. LLMs can be trained to identify pragmatic patterns that are not defined by lexical or syntactic aspects by incorporating corpora marked with speech acts, implicatures, and politeness strategies (Bunt et al., 2010; Danescu-Niculescu-Mizil et al., 2013). This necessitates joint annotation, especially when it comes to languages that have received little research.

b) Multilingual Comparative Fine-Tuning

Multilingual fine-tuning enables models to learn the difference between pragmatic norms across different languages. As an example, parallel corpora may be employed to show how the indirect requests in English are translated to a more direct one in Russian. Such a comparative methodology does not allow models to generalize pragmatic norms of one language to others.

c) Metrics of Pragmatic Evaluation

Pragmatic competence can be evaluated using new evaluation metrics. According to the methodology, pragmatic accuracy can be measured in terms of Implicature Recognition Rate (IRR) and Politeness Preservation Score (PPS). Such measures should be tested on the human judgments, to give reliability.

d) Interactive and Dynamic Context Modeling

Pragmatics is a dynamic element that occurs in interaction. Thus, the models of the future must include the means of monitoring discourse history and adjusting to the norms of a user. Human feedback reinforcement learning (Ouyang et al., 2022) provides a way, which allows models to optimize pragmatic performance over time.

e) Ethical Protection and Cultural Sensitivity

Any integration system should incorporate a system of ethical protection. This involves biased dataset auditing, representation of annotators of diverse backgrounds, and cultural sensitivity mechanisms. The above safeguards are aimed at making sure that pragmatic competence does not reinforce stereotypes or inequalities accidentally.

4.5 Future Research and Application Implications

The application of pragmatics to LLMs has extensive implications. Pragmatically enriched models may keep cultural appropriateness and cultural misconduct may be minimized in translation since they might contain the semantic meaning besides the cultural appropriateness. Such models would be helpful in the field of education, as they can help language learners emphasize pragmatic differences between languages. They might help make the intercultural interactions in business and diplomacy easier, reducing the chances of an accidental insult.

In the research sense, the work has been able to fill the gap between the computational practice and the linguistic theory. It shows how the information about Gricean maxims, Speech Act Theory, and intercultural pragmatics could be implemented in machine learning models. In addition, it emphasizes the necessity of interdisciplinary cooperation: linguists, computer scientists, and cultural professionals have to cooperate to produce pragmatically competent AI.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

5. Conclusion and Future Directions

This paper has discussed the importance of pragmatics in augmenting communicative ability of Large Language Models (LLMs), especially when used in multilingual and cross-cultural environments. Although such LLMs as GPT-4, PaLM, and LLaMA have made a breakthrough in the field of natural language processing with state-of-the-art results in any task, including translation or dialogue generation, their shortcomings in pragmatic reasoning cannot be ignored (Huang et al., 2023; Zhou et al., 2023). As a dimension of language, pragmatics, which deals with meaning based on context, intent, and cultural norms, is poorly covered in the existing models.

The analysis of the available literature illustrated the inference and situational centrality of communication that is underlined in the foundational theories, including Cooperative Principle of Grice (1975), Speech Act Theory (Austin, 1962), and Relevance Theory (Sperber and Wilson, 1986/1995). However, such subtleties are frequently misrepresented by LLMs, and their outputs are semantically correct and pragmatically offensive. Such a limitation is especially problematic in the context of multilingual communication, where the pragmatic norms differ between cultures. The use of politeness in English, directness in Russian, honorifics in Japanese, and formulaic religious language in Arabic all depict the extent to which pragmatic meaning is highly institutionalized in the sociocultural practices (House, 2009; Wierzbicka, 2003; Ide, 1989).

The proposed methodology of the research aimed to overcome these problems by applying pragmatically annotated corpora, experimental fine-tuning, and creation of new evaluation metrics. The strategy shows a way of the path to building pragmatic reasoning into LLMs through combining qualitative analysis with quantitative performance metrics. Notably, the ethical aspect of such a methodology is also anticipated, such as cultural sensitivity, mitigation of bias, and transparency, so that pragmatic enrichment would be equitably offered to a variety of communities (Floridi and Cowls, 2019; Bender et al., 2021).

This discussion highlighted four contributions. First, to be able to approximate human communicative competence, LLLMs must be pragmatically enriched, and not only enriched. Second, the multifaceted nature of multilingual pragmatics underscores the fact that pragmatic transfer errors may occur and that pragmatic fine-tuning should be done comparatively. Third, ethical issues demonstrate that pragmatic norms show social hierarchies and thus, they should be integrated in a considerate way to prevent the possibility of promoting inequalities. Lastly, real world models such as annotated training to practical measure of evaluation have practical avenues of integration.

5.1 Future Directions

In the future, there are a number of directions that should be explored more:

1. Growth to Low-Resource Languages.

   A lot of the existing studies revolve around the high-resource languages like the English and Japanese. Pragmatic enrichment ought to be applied to low-resource languages in the future so that AI systems do not maintain linguistic inequalities in the world (Bender et al., 2021).

2. Adaptation Dynamic Pragmatic Adaptation

   Pragmatic norms do not stay the same but they change with the change of generations, the digital communication practice, and the intercultural interaction (Taguchi, 2019). The

next generation models must be able to make dynamic changes to adapt them so that they can make adjustments to pragmatic competence on the fly based on what the user is saying or the situation on the ground.

3. Multimodal Pragmatic Integration.

Nonverbal communication can be a way of stating pragmatic meaning: the presenter may use changes in tone, gesture, or even facial expression. The addition of multimodal inputs may help LLams to better approximate the way humans reason pragmatically (Li et al., 2021).

4. Ethical Pragmatic AI Governance.

Since pragmatic competence will be a part of AI, governance structures will have to deal with the problem of bias, cultural representation, and accountability. It will be necessary to have collaboration between linguists, ethicists and technologists to assure responsible deployment.

5. Real-World Domain Applications.

Lastly, pragmatically enriched LLMs can potentially be used in a transformative manner in diplomacy, healthcare, education, and business. Application of the models in domain-specific settings will yield some information on the strengths and weaknesses of the models.

5.2 Closing Reflection

To sum up, the inclusion of pragmatics in LLMs is an essential move toward developing AI-based systems that would not only be linguistically correct but also contextually and culturally sensitive. This study helps to develop the vision of language technologies that support the authentic and inclusive communication process by applying a multilingual standpoint. The next step has to involve interdisciplinary effort and long-term dedication to ethical design. In case of realization, pragmatically competent LLMs can revolutionize the world of communication and AI will not be a language processing tool but a cross-cultural understanding companion.

References

Adiwardana, D., Luong, M., So, D., Hall, J., Fiedel, N., Thoppilan, R., … & Le, Q. (2020). *Towards a human-like open-domain chatbot*. arXiv preprint arXiv:2001.09977.

Al-Khatib, M. A. (2001). The pragmatics of invitation making and acceptance in Jordanian society. *Journal of Pragmatics, 33*(10), 1905–1928.

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).

Blum-Kulka, S. (1989). Playing it safe: The role of conventionality in indirectness. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 37–70). Ablex.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., … & Amodei, D. (2020). *Language models are few-shot learners*. arXiv preprint arXiv:2005.14165.

Bunt, H., Alexandersson, J., Carletta, J., Choe, J.-W., Fang, A. C., Hasida, K., … & Zimmermann, D. (2010). Towards an ISO standard for dialogue act annotation. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)* (pp. 2548–2555).

Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., … & Fiedel, N. (2022). *PaLM: Scaling language modeling with pathways*. arXiv preprint arXiv:2204.02311.

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems, 30*.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., & Potts, C. (2013). A computational approach to politeness with application to social factors. In *Proceedings of ACL 2013* (pp. 250–259).

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT* (pp. 4171–4186).

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1).

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts* (pp. 41–58). Academic Press.

Haugh, M. (2018). The metapragmatics of culture. In I. Kecskes (Ed.), *Exploring intercultural communication: Language in action* (pp. 33–52). Routledge.

Haugh, M., & Chang, W. (2019). Understanding sarcasm across cultures: The role of context and prosody. *Journal of Pragmatics, 147*, 1–15.

House, J. (2009). *Cross-cultural pragmatics: Theory and practice*. Oxford University Press.

Huang, T., Lin, Z., & Yu, Z. (2023). Benchmarking large language models on pragmatic tasks. In *Proceedings of ACL 2023* (pp. 555–567).

Ide, S. (1989). Formal forms and discernment: Two neglected aspects of universals of linguistic politeness. *Multilingua, 8*(2–3), 223–248.

Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Routledge.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., … & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), 1–38.

Kecskes, I. (2014). *Intercultural pragmatics*. Oxford University Press.

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Lewis, M., Perez, E., Piktus, A., Karpukhin, V., Goyal, N., Küttler, H., … & Zettlemoyer, L. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems, 33*.

Li, X., Su, S., & Zhao, T. (2021). Multimodal pragmatics: Towards context-aware dialogue systems. *Computational Linguistics, 47*(4), 831–857.

Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. In *Text summarization branches out: Proceedings of the ACL-04 Workshop* (pp. 74–81).

Mills, S. (2003). *Gender and politeness*. Cambridge University Press.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., … & Christiano, P. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*.

ISSN E: 2709-8273
aa ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proceedings of ACL 2002* (pp. 311–318).

Roller, S., Dinan, E., Goyal, N., Ju, D., Williamson, M., Liu, Y., … & Weston, J. (2021). Recipes for building an open-domain chatbot. In *Proceedings of EACL 2021* (pp. 300–325).

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell. (Original work published 1986)

Taguchi, N. (2019). *Pragmatics in language teaching and learning*. Routledge.

Thomas, J. (1995). *Meaning in interaction: An introduction to pragmatics*. Routledge.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., … & Lample, G. (2023). *LLaMA: Open and efficient foundation language models*. arXiv preprint arXiv:2302.13971.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems, 30*.

Wierzbicka, A. (2003). *Cross-cultural pragmatics: The semantics of human interaction* (2nd ed.). Mouton de Gruyter.

Zhou, K., Xu, Y., & Li, J. (2023). Beyond semantics: Evaluating pragmatic competence in large language models. *Journal of Artificial Intelligence Research, 77*, 445–472.