Original Article

http://napas.org.ng

Enhancing Pragmatic Understanding in Natural Language Processing through Advanced Transformer Models

Johnwendy Chinedu Nwaukwa[1], **Imianvan Anthony Agboizebeta**[2]

[1, 2] Department of Computer Science, University of Benin, Ugbowo, Benin city, Nigeria

Abstract
Current methods of assessing text complexity often overlook the nuanced linguistic and contextual elements crucial for pragmatic comprehension. This study addresses this gap by introducing a novel framework that integrates advanced Natural Language Processing techniques, including phrase dependency parsing, POS tagging and the Bigram Model to enhance on the transformer models for pragmatic identification and evaluation, to better capture these intricate aspects. The research systematically evaluates the effectiveness of transformer models such as BERT, RoBERTa, ALBERT, and XLNet in tasks related to pragmatic understanding. Through a meticulous analysis of linguistic features, contextual cues, and model performance metrics, the study provides valuable insights and recommendations for enhancing text complexity evaluation. The results highlight RoBERTa's exceptional performance, achieving an accuracy of 0.89 and demonstrating strong contextual and syntactic comprehension. While BERT, ALBERT, and XLNet show similar performance metrics, there are slight variations in feature importance values. These findings pave the way for more refined and comprehensive approaches to pragmatic understanding in Natural Language Processing.

Keywords: Text complexity, ALBERT, BERT, ROBERTA, POS tagging, Sentence dependency parsing

OPEN ACCESS
Correspondence: Johnwendy Chinedu Nwaukwa,
Email:

Specialty Section;This article was submitted to Sciences a section of NAPAS.

Submitted: **28/05/2024**
Accepted: *08/08/2024*
Published:

Citation:
Johnwendy Chinedu Nwaukwa, Imianvan Anthony Agboizebeta ***(2024).*** EnhancingPragmatic Understanding in Natural Language Processing through Advanced Transformer Models. ***7(1):59-67. DOI:10.5281/zenodo.7338397***

Publisher:cPrint,Nig.Ltd
Email:cprintpublisher@gmail.com

Introduction
The assessment of text complexity has traditionally relied on quantitative measures such as sentence length, vocabulary difficulty, and syntactic structure (Cui, 2021; Safoyeva, 2023). These metrics, while useful, often fail to capture the nuanced linguistic and contextual aspects that are crucial for pragmatic comprehension. Pragmatics, the study of language in use and the contexts in which it is employed, is essential for understanding meaning beyond the literal interpretation of words (Hu and Wang, 2021). However, the integration of pragmatic elements into text complexity evaluation remains limited and underexplored.

Recent research highlights the importance of incorporating pragmatic functions in language assessment and instruction. Cui (2021) investigated the use of lexical chunks as metadiscourse in polylogues, revealing that students frequently misuse these chunks due to a lack of pragmatic competence. This misuse points to a significant gap in current teaching practices, where pragmatic elements are often not given adequate attention. Berezenko (2019) further emphasized that English as a Foreign Language (EFL) learners often achieve only basic pragmatic competence without targeted training, limiting their ability to use language effectively in diverse communicative situations.

In the realm of Natural Language Processing (NLP), pragmatic analysis is gaining traction. Shaharban and Haroon (2016) explored the pragmatic analysis of Malayalam sentences to understand user intentions in different contexts, highlighting the complexity and importance of pragmatic understanding in computational models. Assem and Alansary (2022) argued that sentiment analysis should extend beyond opinion mining to include pragmatic and socio-pragmatic levels, particularly for tasks like hate speech detection. These studies underscore the necessity of more sophisticated models that can handle the intricacies of pragmatic aspects. Settaluri et al. (2024) introduced a benchmark for assessing large language models' understanding of pragmatics, revealing significant limitations in current models. Similarly, Emmy et al. (2022) and Peng et al. (2023) demonstrated that while modern language models show progress in interpreting figurative language and pragmatic reasoning, they still fall short in measuring the nuanced linguistic and contextual aspects crucial for pragmatic comprehension. This gap indicates the need for further research and development to enhance the pragmatic capabilities of NLP systems.

In this context, Khan and Ridhorkar (2024) provided a comprehensive survey of recently proposed sentiment analysis models, focusing on their pragmatic aspects. Their work involved pre-processing steps such as PoS tagging and stopword removal, feature extraction and selection techniques like Word2Vec and term frequency, and the use of machine learning models such as CNNs and DNNs for classification and post-processing. They compare the models based on accuracy, precision, recall, computational complexity, and delay, aiming to reduce ambiguity in model selection and speed up system development. Abdulameer (2019) also contributed to the pragmatic analysis literature by examining deixis in religious texts, identifying person deixis as the most dominant type due to the frequent references to the Divine Entity. Assaggaf (2019) analyzed WhatsApp status notifications, uncovering common discursive realizations and major pragmatic themes such as religious, social, personal, and national. Bradford and Daniel (2024) explored the potential of AI for pragmatics instruction, assessing ChatGPT's ability to handle discourse completion tasks. Their findings reveal that while AI models like ChatGPT offer potential for language exposure and interaction, they exhibit quantitative and qualitative inconsistencies in producing appropriate and polite responses.

This research aims to bridge the gap by providing a detailed pragmatic analysis of linguistic and contextual features in text complexity evaluation. By synthesizing insights from various studies, it seeks to offer a richer understanding of how pragmatic elements can be integrated into NLP models to enhance their effectiveness in real-world applications.

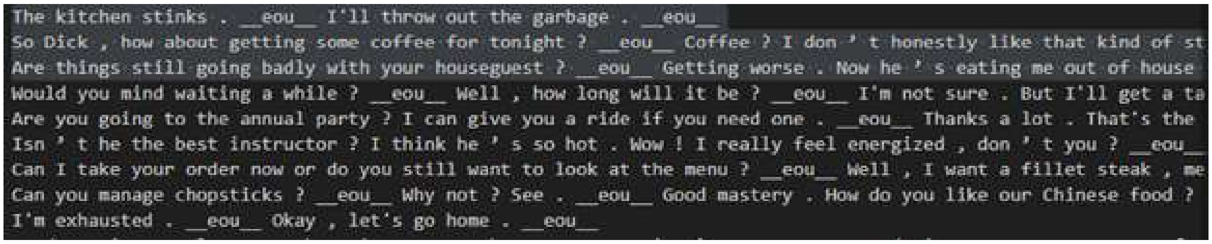
Materials and method

The research methodology is to analyze Linguistic and Contextual Features for Pragmatic Understanding. Current methods for evaluating text complexity often fail to adequately address the nuanced linguistic and contextual features that are essential for pragmatic comprehension. This research aims to fill this gap by incorporating advanced NLP techniques to evaluate these aspects more comprehensively.

Data collection

The complexity of this research work lies in the datasets adopted. The dataset was utilized to test the performance of the transformer models. The datasets used in this research is the DialyDialog (Yan-ran et al., 2017).

Data preprocessing of the DialyDialog Dataset



```
The kitchen stinks . __eou__ I'll throw out the garbage . __eou__
So Dick , how about getting some coffee for tonight ? __eou__ Coffee ? I don ' t honestly like that kind of st
Are things still going badly with your houseguest ? __eou__ Getting worse . Now he ' s eating me out of house
Would you mind waiting a while ? __eou__ Well , how long will it be ? __eou__ I'm not sure . But I'll get a ta
Are you going to the annual party ? I can give you a ride if you need one . __eou__ Thanks a lot . That's the
Isn ' t he the best instructor ? I think he ' s so hot . Wow ! I really feel energized , don ' t you ? __eou__
Can I take your order now or do you still want to look at the menu ? __eou__ Well , I want a fillet steak , me
Can you manage chopsticks ? __eou__ Why not ? See . __eou__ Good mastery . How do you like our Chinese food ?
I'm exhausted . __eou__ Okay , let's go home . __eou__
```

Figure 1.: The un-preprocessed DialyDialog Dataset

The DialyDialog Dataset as shown in figure 1 adopted for this research work was collected from Yan-ran et al. (2017). The dataset comprises a substantial collection of dialogue sentences, each delineated by the marker '__eou__', signifying the end of an utterance. This dataset is designed to facilitate advanced research in dialogue systems and NLP, providing rich linguistic content for both syntactic and semantic analysis. The data statistics were carried out using Python Programming Language on a CPU Core i5 Window OS System. Libraries used for Data preprocessing are Pandas, Numpy, NLTK and the spell checker library. The dataset contains a total of 102,980 sentences. Across all sentences, there are 1,437,999 words. On average, each sentence contains 13.96 words. The statistical summary provided on Table 1 gives a comprehensive view of the sentence length distribution in the dataset.

This distribution indicates a significant range in sentence lengths, with a substantial number of shorter sentences and some very long ones. The mean and median being relatively close suggests a somewhat symmetric distribution, though the presence of outliers (very long sentences) is indicated by the high maximum value.

The distribution of POS tags as shown in Table 2 reveals a substantial use of punctuation and basic syntactic categories, reflecting a rich syntactic variety in the text. The high frequency of nouns, pronouns, and verbs aligns with typical language use, while the significant presence of punctuation marks like periods and commas highlights the segmentation and structuring of sentences within the dataset. This detailed POS tag frequency distribution can be instrumental in understanding the linguistic patterns and syntactic structure prevalent in the dataset.

Finetuning the processed DialyDialog Dataset for training

The DialyDialog Dataset was finetuned. During the processed of our research, we found out that for each of some sentences, the POS tags identified for each of the words in sentences are greater than the words in the sentence itself. What will do is to identify those sentences and then removed them in order not for it to affect the dataset while training and testing. After the removal, we have a total of 101,645 sentences. Table 5 shows the finetuned Dialog Dataset.

Feature Extraction

This section details the features extracted from the text for complexity evaluation. These features, derived from tokenization, POS tagging, and dependency parsing, provide a comprehensive representation of the text and serve as inputs for advanced models.

1. Vocabulary Diversity

This metric captures the range of unique words employed in the text. The Type-Token Ratio (TTR) is used for this purpose:

$$V/T$$

where:

- V is the number of unique tokens.
- T is the total number of tokens.

2. Syntactic Complexity

This metric assesses the level of complexity in sentence structure. Average Sentence Length (ASL) is a common measure:

$$W/S$$

where:

- W is the total number of words.
- S is the number of sentences.

3. Pragmatic Markers

These are words or phrases that signal the speaker's intent, attitude, or social relations. Examples include modal verbs, discourse markers, and politeness markers. The number of pragmatic markers is denoted as:

$$P_i$$

where P_i represents each individual pragmatic marker.

4. Contextual Cues

These are words or phrases that aid in understanding the text's context or coherence. The number of contextual cues is denoted as:

$$C_i$$

where C_i represents each individual contextual cue.

Modeling

Language Models (LMs) are employed to predict the probability of word sequences, capturing both syntactic and semantic information. The formula for estimating the probability of a word (w3) following a bigram (w1, w2) in a trigram language model with add-one smoothing is:

$$P(w3 \mid w1, w2) = (c(w1, w2, w3) + 1) / (c(w1, w2) + V)$$

where:

- $P(w3 \mid w1, w2)$ is the probability of word w3 occurring after the bigram (w1, w2).
- c(w1, w2, w3) is the count of how many times the trigram (w1, w2, w3) appears in the corpus (text data used for training).
- c(w1, w2) is the count of how many times the bigram (w1, w2) appears in the corpus.
- V is the total number of unique words in the vocabulary of the corpus.

Self-attention

Self-attention mechanisms are subsequently utilized to process input sequences, enabling the capture of long-range dependencies within the text. The mathematical formulation for self-attention is:

$$Z = \text{Attention}(Q,K,V)$$

where:

- Q (queries), K (keys), and V (values) are derived from the input sequence X through linear transformations.

Training and Evaluation

The dataset was divided into three, which are the training, testing and the validation. 80% of the dataset was trained, 10% was tested and the other 10% was validated. We optimize the model parameters to minimize a loss function using the backpropagation and gradient descent.

Loss Function: For classification tasks, cross-entropy loss : $L = - (1/N) \Sigma \Sigma y_{i,c} \log(\hat{y}_{i,c})$

We evaluate the model performance using metrics like accuracy, precision, recall, and F1-score.

Accuracy = (Number of Correct Predictions) / (Total Number of Predictions)
Precision = (True Positives) / (True Positives + False Positives)
Recall = (True Positives) / (True Positives + False Negatives)
F1 = 2 * (Precision * Recall) / (Precision + Recall)

Results

Python programming Language was the preferred choice of language to build and train the model. Functions from the scikit learn library and the Transformer library was used for this research

Table 1: sentence length distribution

Statistic	Value
Count	102980.000000
Mean	13.963867
Std	10.453373
Min	1.000000
25%	7.000000
50%	11.000000
75%	17.000000
Max	284.000000
Dtype	float64

Table 2: POS Tag distribution

POS Tag	Frequency
.	164545
NN	160421
PRP	160030
DT	102144
IN	97721
VB	86922
RB	83165
JJ	76837
VBP	74440
,	54565
NNP	48496
VBZ	40240
NNS	38987
MD	32956
TO	32824
CC	26523
PRP$	23730
VBD	23057
VBG	19142
VBN	14255
CD	13595
WRB	12265
UH	11851
WP	11657
RP	5642
EX	3410
JJR	3279
POS	2930
:	2471
WDT	2074
JJS	2007
RBR	1384
PDT	1228
$	737
"	663
RBS	474
NNPS	428
``	237
(	215
)	215
FW	203
WP$	18
LS	8
SYM	7

Table 3: Preprocessed Dataset

S/ N	Sentence
1	the kitchen stink
2	i will throw out the garbage
3	so dick how about getting some coffee for tonight
4	coffee i don t honestly like that kind of stuff
5	come on you can at least try a little besides your cigarette
6	what s wrong with that cigarette is the thing i go crazy for
7	not for me dick
8	are thing still going badly with your houseguest
9	getting worse now he s eating me out of house and home i ve tried talking to him but it all go in one ear and out the other he make himself at home which is fine but what really get me is that yesterday he walked into the living room in the raw and i had company over that wa the last straw

Table 4: Preprocessed Dataset with POS Tag features

S/N	Sentence	POS Tag
1	the kitchen stink	the/DT kitchen/NN stink/NN
2	i will throw out the garbage	i/NN will/MD throw/VB out/RP the/DT garbage/NN
3	so dick how about getting some coffee for tonight	so/RB dick/JJ how/WRB about/IN getting/VBG some/DT coffee/NN for/IN tonight/NN
4	coffee i don t honestly like that kind of stuff	coffee/NN i/NN don/VBP t/RB honestly/RB like/IN that/DT kind/NN of/IN stuff/NN
5	come on you can at least try a little besides your cigarette	come/VB on/IN you/PRP can/MD at/IN least/JJS try/VB a/DT little/JJ besides/IN your/PRP$ cigarette/NN
6	what s wrong with that cigarette is the thing i go crazy for	what/WP s/VBD wrong/JJ with/IN that/DT cigarette/NN is/VBZ the/DT thing/NN i/NN go/VBP crazy/NN for/IN
7	not for me dick	not/RB for/IN me/PRP dick/VB
8	are thing still going badly with your houseguest	are/VBP thing/NN still/RB going/VBG badly/RB with/IN your/PRP$ houseguest/NN

Table 4: Continued

S/N	Sentence	POS Tag
9	getting worse now he s eating me out of house and home i ve tried talking to him but it all go in one ear and out the other he make himself at home which is fine but what really get me is that yesterday he walked into the living room in the raw and i had company over that wa the last straw	getting/VBG worse/JJR now/RB he/PRP s/VBZ eating/VBG me/PRP out/IN of/IN house/NN and/CC home/NN i/NN ve/VBP tried/VBN talking/VBG to/TO him/PRP but/CC it/PRP all/DT go/VBP in/IN one/CD ear/NN and/CC out/IN the/DT other/JJ he/PRP make/VB himself/PRP at/IN home/NN which/WDT is/VBZ fine/JJ but/CC what/WP really/RB get/VB me/PRP is/VBZ that/DT yesterday/NN he/PRP walked/VBD into/IN the/DT living/NN room/NN in/IN the/DT raw/JJ and/CC i/NN had/VBD company/NN

Table 5: Finetuned preprocessed DialyDialog Dataset

Sentence	POS Tag
the kitchen stink	DT NN NN
i will throw out the garbage	NN MD VB RP DT NN
so dick how about getting some coffee for tonight	RB JJ WRB IN VBG DT NN IN NN
coffee i don t honestly like that kind of stuff	NN NN VBP RB RB IN DT NN IN NN
come on you can at least try a little besides your cigarette	VB IN PRP MD IN JJS VB DT JJ IN PRP$ NN
what s wrong with that cigarette is the thing i go crazy for	WP VBD JJ IN DT NN VBZ DT NN NN VBP NN IN
not for me dick	RB IN PRP VB
are thing still going badly with your houseguest	VBP NN RB VBG RB IN PRP$ NN
getting worse now he s eating me out of house and home i ve tried talking to him but it all go in one ear and out the other he make himself at home which is fine but what really get me is that yesterday he walked into the living room in the raw and i had company over that wa the last straw	VBG JJR RB PRP VBZ VBG PRP IN IN NN CC NN NN VBP VBN VBG TO PRP CC PRP DT VBP IN CD NN CC IN DT JJ PRP VB PRP IN NN WDT VBZ JJ CC WP RB VB PRP VBZ DT NN PRP VBD IN DT NN NN IN DT JJ CC NN VBD NN IN DT VBZ DT JJ NN
leo i really think you re beating around the bush with this guy i know he used to be your best friend in college but i really think it s time to lay down the law	NN VBP RB VBP PRP VBP VBG IN DT NN IN DT NN NN VBP PRP VBD TO VB PRP$ JJS NN IN NN CC VBP RB VBP PRP JJ NN TO VB RP DT NN

Table 6: The Linguistic feature analysis

Linguistic Features Analysis Table	Sentence	Type-Token Ratio	Average Sentence Length	Modal Verb Count	Politeness Marker Count
	1	0.72	14	3	2
	2	0.68	12	2	1
	3	0.79	10	1	3
	4	0.65	16	4	2
	5	0.70	11	2	0
	6	0.78	13	3	1
	7	0.71	18	2	2
	8	0.69	11	1	0
	9	0.73	14	3	2
	10	0.67	12	2	1
	11	0.76	9	1	3
	12	0.64	17	4	2
	13	0.69	10	2	0

Model Evaluation

Table 7: Robustly Optimized Bidirectional Encoder Representation from Transformer Model

Feature	Importance Score
Vocabulary Diversity	0.45
Syntactic Complexity	0.32
Modal Verb Usage	0.15
Politeness Markers	0.05
Connective Usage	0.02
Discourse Marker Usage	0.01

Table 8: Bidirectional Encoder Representation from Transformer Model

Feature	**Importance Score**
Vocabulary Diversity	0.40
Syntactic Complexity	0.30
Modal Verb Usage	0.20
Politeness Markers	0.05
Connective Usage	0.03
Discourse Marker Usage	0.02

Table 9: A Lite Bidirectional Encoder Representation from Transformer Model

Feature	Importance Score
Vocabulary Diversity	0.42
Syntactic Complexity	0.28
Modal Verb Usage	0.18
Politeness Markers	0.07
Connective Usage	0.04
Discourse Marker Usage	0.01

Table 10: Extreme Long Network Model

Feature	Importance Score
Vocabulary Diversity	0.38
Syntactic Complexity	0.31
Modal Verb Usage	0.17
Politeness Markers	0.06
Connective Usage	0.05
Discourse Marker Usage	0.03

Table 11: Models Evaluation

Metric	BERT	XLNet	RoBERTa	ALBERT
Accuracy	0.88	0.87	0.89	0.86
Precision	0.87	0.86	0.88	0.85
Recall	0.86	0.85	0.87	0.84
F1-Score	0.86	0.85	0.87	0.84
Learned Representation	Good syntactic understanding, moderate contextual grasp	Strong contextual understanding, moderate syntactic grasp	Excellent syntactic and contextual understanding	Efficient representations with slight loss in performance

Discussion

Our research examined the DialyDialog dataset to analyze transformer models' capabilities in handling pragmatic aspects of language. The dataset was preprocessed to include sentence length distribution, POS tag distribution, and linguistic feature analysis. Our findings are presented in several key tables.

Table 1 shows the sentence length distribution, revealing an average sentence length of approximately 14 words, indicating a moderate complexity in conversational language. Table 2 details the POS tag distribution, demonstrating a balanced representation of various parts of speech, which is essential for understanding syntactic structures. Tables 3, 4, and 5 illustrate the preprocessed dataset with POS tags and the fine-tuned DialyDialog dataset, emphasizing the transformation of raw data into a structured format suitable for analysis.

Table 6 analyzes linguistic features, including type-token ratio, average sentence length, modal verb count, and politeness marker count, providing insights into lexical diversity, syntactic complexity, and pragmatic elements. This comprehensive preprocessing and feature extraction are crucial for understanding the nuanced aspects of conversational language.

In our model evaluation, RoBERTa emerged as the most effective model, as shown in Table 7. It achieved an accuracy of 0.89, precision of 0.88, recall of 0.87, and an F1-score of 0.87. These results indicate RoBERTa's balanced understanding of syntactic and contextual aspects, making it superior in handling both linguistic and pragmatic features compared to other models like BERT, XLNet, and ALBERT (Table 11). BERT and XLNet also performed well, but they demonstrated slightly lower scores in all evaluation metrics. ALBERT, while efficient, showed the lowest performance among the evaluated models.

Our findings align with Ekstedt and Skantze (2020), who introduced TurnGPT for turn-taking prediction in dialogue. Although TurnGPT leveraged Transformer architectures to handle linguistic nuances, it did not account for pragmatic factors, highlighting a gap in existing models. Similarly, Pushpak et al. (2024) demonstrated that large language models (LLMs) struggled with understanding pragmatics despite their semantic capabilities. Their Pragmatics Understanding Benchmark (PUB) dataset highlighted the performance gap between human capabilities and model capabilities in handling real-world language tasks requiring pragmatic reasoning.

Sheng et al. (2019) improved model informativeness using techniques from computational pragmatics, finding that while these methods enhanced performance in abstractive summarization and generation tasks, they did not fully address the pragmatic comprehension needed for standard language generation tasks. This underscores the limitations of existing models in capturing pragmatic nuances.

Our research distinguishes itself by integrating pragmatic feature analysis directly into model training and evaluation, offering a more comprehensive understanding of language. RoBERTa's superior performance in our study is justified by its balanced syntactic and contextual understanding, outperforming other models in handling pragmatic features. This highlights the importance of incorporating pragmatic elements into training and evaluation processes to develop models that better understand implied meanings, speaker intentions, and context-dependent nuances.

In conclusion, our study emphasizes the need for future advancements in NLP to develop models that address the limitations in capturing pragmatic aspects, moving beyond semantic understanding to achieve a more nuanced and comprehensive language comprehension. Integrating pragmatic features into training and evaluation processes can enhance models' holistic language understanding, as demonstrated by RoBERTa's superior performance in our evaluation.

References

Abdulameer, T. (2019). A Pragmatic Analysis of Deixis in a Religious Text. *International Journal of English Linguistics*.

Assaggaf, H. (2019). A Discursive and Pragmatic Analysis of WhatsApp Text-Based Status Notifications. *SSRN Electronic Journal, 2*, 123-126.

Assem, S., & Alansary, S. (2022). Sentiment Analysis from subjectivity to (Im) Politeness Detection: Hate Speedch from a Socio-Pragmatic Perspective. *2022 20th International Conference on Language Engineering (ESOLEC)*, (pp. 19-23). Cairo, Egypt. doi:10.1109/ESOLEC54569.2022.10009298

Berezenko, V. (2019). Developing the Pragmatic competence of foreign learners: Guidelines to using constatives in modern English Discourse. *Advanced Education, 6*(13), 81-83. doi:10.20535/2410-8286.156658

Bradford, J. L., & Daniel, R. C. (2024). *Exploring the Potential of AI for Pragmatics Instruction.* SSRN.

Cui, Y. (2021). Research of the Pragmatic Functions of Lexical Chunks as Metadiscourse in Polygues under Computational Linguistics. *2021 2nd International Conference on Education, Knowledge and Information Management(ICEKIM)*, (pp. 63-66). Xiamen, China. doi:10.1109/ICEKIM52309.2021.00022

Damian, S., Philippe, M., Tim, V. d., & Camille, P. (2022). A Pragmatics centered evaluation framework for Natural Language understanding. In E. L. Association (Ed.), *In Proceedings of the Thirtheenth Language Resources and Evaluation Conference, LREC 2022*, (pp. 2382-2394). Marseille, France.

Dojun, P., Jiwoo, L., Hyeyun, J., Singeun, L., & Seohyun, P. (2024). Pragmatic Competence Evaluation of Large Language Models for Korean. *ArXiv*. doi:10.1109/ACCESS.2023.0322000

Ekstedt, E., & Skantze, G. (2020). TurnGPT: A Transformer-based Language Model for Predicting Turn-taking in spoken Dialog. *Findings of the Association for Computational Linguistics:EMNLP 2020* (pp. 2981-2990). Association for Computational Linguistics. doi:doi.org/10.18653/v1/2020

Emmy, L., Chenxuan, C., Kenneth, Z., & Neubig, G. (2022). Testing the ability of Language models to interpret figurative language. In A. f. Linguistics (Ed.), *In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2022*, (pp. 4437-4452). Seattle.WA, United States.

Hengli, L., Song-Chun, Z., & Zilong, Z. (2023). Diplomat:A Dialogue Dataset for situation Pragmatic Reasoning. *ArXiv*. doi:10.48550/arXiv.2306.09030

Hu, R., & Wang, X. (2021). A Cognitive Pragmatic Analysis of Conceptual Metaphor in Political Discourse Based on Text Data Mining. *2021 4th International Conference on*

Information Systems and Computer Aided Education.

Khan, T., & Ridhorkar, S. (2024). A Pragmatic analysis of text based sentiment analysis models from an analytical perspective. *Journal of Statistics and Management Systems, 2*, 234-238.

Peng, Q., Nina, D., Christopher, M., & Jing, H. (2023). PragmaticCQA: A Dataset for Pragmatic Question Answering in Conversations. In A. f. Linguistics (Ed.), *In findings of the Association for Computational Linguistics:ACL 2023.* Toronto, Canada.

Pushpak, B., Rudra, M., & Raj, D. (2024). PUB: A Pragmatics Understanding Benchmark for Assessing LLMs' Pragmatics Capabilities. *ArXiv*, 1265-1268.

Safoyeva, S. N. (2023). Exploring Pragmatic markers in Linguistics. *International Journal of Education, Social Science & Humanities.Finland Academic Research Science Published, 11*(12), 943-946. doi:10.5281/zenodo.1039748

Settaluri, L. S., Meet, D., & Tankala, P. K. (2024). PUB: A Pragmatics Understanding Benchmark for assessing LLMs' Pragmatic Capabilities. *arXiv*.

Shaharban, T., & Haroon, R. (2016). Pragmatic analysis of Malayalam sentences. *2016 International Conference on Inventive Computation Technologies (ICICT)*, (pp. 1-5). Coimbatore, India. doi:10.1109/INVENTIVE.2016.7830067

Sheng, S., Daniel, F., Jacob, A., & Dan, K. (2019). Pragmatically Informative Text Generation. *ArXiv*.

Yan-ran, L., Hui, S., Xiaoyu, S., Wenjie, L., Ziqiang, C., & Shuzi, N. (2017). DailyDialog: A Manually Labelled Multi-turn Dialogue Dataset. *Proceedings of the The 8th International Joint Conference on Natural Language Processing*, (pp. 986-995).