


# ChatGPT Is Powerful, but Does It Have Power Distance?

## A Study of Culturally Imbued Discourse in AI-Generated Essays

Andrew Schenck, Fort Hays State University, USA\*

 <https://orcid.org/0000-0002-3864-6267>

### ABSTRACT

Power distance (PD), a cultural value denoting acceptance of asymmetrical power relationships, influences the force of rhetoric used by a writer to address their reader. However, AI technologies such as ChatGPT lack an explicit awareness of PD, which could affect the quality of AI-generated persuasive texts used for language learning. To investigate this issue, 200 persuasive essays written by ChatGPT were compared to 200 essays written by L1-English university learners. Three elements of formulaic language related to PD were examined: stances, modals, and pronoun deixis. Differences in stances ( $z = -3.411$ ;  $p = .001$ ) and modals ( $z = -2.100$ ;  $p = .036$ ) were both significant according to the Wilcoxon signed ranks formula, whereas differences in pronoun deixis were nearly significant ( $z = -1.917$ ;  $p = .055$ ). Overall, language of ChatGPT appears generic and incomplete, suggesting that consistent and uniform expressions are being borrowed from an LLM training corpus to mimic aspects of PD. Limitations of AI highlight a need for pedagogical emphasis of culturally imbued discourse.

### KEYWORDS

ChatGPT, Corpus, Discourse, English, LLM, Modals, Power Distance, Pronoun Deixis, Stances, Writing

### INTRODUCTION

Social media outlets like Twitter and TikTok have recently been abuzz with posts about AI platforms like ChatGPT (Haensch et al., 2023; Li et al., 2023). Despite concerns expressed by some educators concerning possible violations of academic integrity and adverse consequences to learning outcomes (Li et al., 2023; Yan, 2023), attitudes about AI technology remain largely positive (Haensch et al., 2023; Liu & Ma, 2023). The potential of AI to enhance English instruction has also been acknowledged (Jeon & Lee, 2023; Mohamed, 2023). ChatGPT may serve as an interlocutor, teaching assistant, content provider, or evaluator (Jeon & Lee, 2023). It may also allow students to create texts from diverse perspectives, enhancing inductive forms of learning.

LLMs may have a plethora of applications in writing instruction. Recent research suggests that ChatGPT is a valuable tool for correcting surface-level errors related to grammar or structure

DOI: 10.4018/IJAET.338219

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

(Algaraady & Mahyoob, 2023). The Modern Language Association and the National Council of Teachers of English also point out that teachers may use LLMs in the following ways to support aspects of style and rhetoric:

- To enhance students' rhetorical knowledge, critical thinking, and knowledge of conventions.
- To offer a practical demonstration of some critical rhetorical concepts that have influenced writing and rhetoric studies, especially related to questions of process, praxis, and the construction of meaning.
- To provide modals of written prose that can be used to highlight differences in genre, tone, diction, literary style, and disciplinary focus. Teachers can use LLMs to offer new processes for students developing multimodal writing genres since LLMs can process multimodal inputs and outputs. (MLA-CCCC Joint Task Force on Writing and AI, 2023, pp. 8-9)

As suggested above, LLMs like ChatGPT have the potential to help students learn more sophisticated techniques for constructing rhetoric, which may lead to responses that vary in genre, grammar, and style. However, little research has been conducted to ensure that LLMs live up to these lofty expectations. More research is needed to understand how teachers may effectively use AI technology to promote better writing (Jeon & Lee, 2023; Mohamed, 2023). Albeit limited, some research suggests that ChatGPT cannot adequately construct or interpret complex rhetoric, which may hamper efforts to integrate technology in English writing classrooms (Algaraady & Mahyoob, 2023; Fan & Jiang, 2023). These preliminary findings are intriguing, yet more research is needed to concretely identify and address the limitations of AI technology, thereby ensuring that the benefits of LLMs like ChatGPT can be fully realized.

## LITERATURE REVIEW

While LLMs can potentially transform teaching and assessment of writing, the complexity of human discourse or rhetoric cannot easily be replicated by AI technologies. Using an essay prompt that examined the difference between corpus linguistics and discourse analysis, a professor from the University of Murcia put ChatGPT to the test, finding that AI-generated essays were well structured and showed a cogent argument yet lacked clear appeals to authority (Pérez-Paredes, 2023). Inadequacies of ChatGPT were further exposed by Ian Bogost (2022) in his article entitled "ChatGPT is Dumber than You Think"; he found the enthusiasm for ChatGPT to be "misplaced," writing that:

*ChatGPT cannot truly understand the complexity of human language and conversation. It is simply trained to generate words based on a given input, but it cannot truly comprehend the meaning behind those words. This means that any responses it generates are likely to be shallow and lacking in depth and insight. (para. 2)*

Because LLMs rely on frequent linguistic structures to replicate a human's writing, they may not have the complex understanding of human nature needed to comprehend or replicate discourse. Although it has become apparent that AI-generated texts lack a human touch, the degree to which LLMs like ChatGPT can or cannot use aspects of human discourse is not well known.

If LLMs are to be used in the classroom, educators must understand how elements of natural human discourse are replicated, along with deficiencies that require pedagogical intervention. Effective writing requires more than just a knowledge of structure or morphosyntax, which may put ChatGPT at a distinct disadvantage. According to Hyland (2018), a knowledge of meta-discourse is required. Meta-discourse is defined as:

*The cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community. (Hyland, 2018, p. 43)*

A key term in this definition is “interactional meanings,” which signifies a complex understanding of the writer and his or her relationship with the reader. Just as professors alter language to interact politely with their students, writers modify persuasive rhetoric by readers in specific contexts or speech communities (Biber, 1996; Halliday & Hasan, 1989). Due to the complexities of meta-discourse, generating or evaluating effective rhetoric may be challenging for an LLM like ChatGPT. More research is required to determine both strengths and weaknesses of such technology. This research may help in the design of better instruction that integrates ChatGPT with other examples of authentic English discourse. Such research may have implications for English first and second-language learning contexts, which both require an understanding of discourse in the L1.

### **Culturally Imbued Language and Limitations of ChatGPT**

LLMs like ChatGPT may have distinct applications in writing pedagogy, yet the degree to which they can utilize discourse for specific tasks or speech communities appears to be limited. This limitation may be better understood by comparing AI-generated texts with writings from specific cultural contexts. National or regional belief systems currently shape human writers’ choices when designing discourse for specific situational contexts (Hamadouche, 2013; Hamam, 2020; Hammad, 2002; Jiang, 2006; Paltridge, 2012). This view is illustrated by a study of five different American universities, which showed that students who learned Chinese struggled to adopt a new rhetorical style referred to as *yilùnwén* writing, choosing to rely on traditional rhetorical patterns used in their L1 (Liu & Du, 2018, p. 1). Such a study reveals that cultural values influence the expression of rhetoric or other forms of discourse in writing. By comparing these values with compositions produced by ChatGPT, new suggestions for integrating AI with writing instruction may be gleaned.

Although cultural values shape discourse, little is known about how such values are, or are not, imbued in the language generated by LLMs like ChatGPT. One cultural value of significance to writing is Power Distance (PD), which denotes the degree to which a cultural group accepts (or does not accept) asymmetrical power relationships (Shah et al., 2014). This value may directly impact a writer’s position of authority over their reader. PD varies considerably in the writings of human populations, with high PD cultures accepting differences in authority (as well as the distance between the writer and the reader) and low PD cultures rejecting unequal power relationships to form potentially closer or less formal bonds. In a study of 344 emails written by 110 students, students from low PD cultures like Norway used less formal discourse than their high PD counterparts from other countries (Bjørge, 2007). High PD has also been found in legal documents related to international trade, which use formulaic language to denote asymmetries between writer and reader (Orts, 2016).

Being designated as low PD countries, Western contexts such as the United States, United Kingdom, Australia, Canada, and New Zealand may utilize discourse that connects more closely with the reader on an egalitarian level (Northouse, 2021). However, little is known about how these elements of PD are used in persuasive writings constructed by ChatGPT, revealing a need for further research. Past research on PD has emphasized persuasive writings by human writers, revealing key insights about differences in how opinions are expressed. In a study by Mattheoudakis and Hatzitheodorou (2011), writers from low PD countries tended to question the authority of experts more often, whereas Greek learners with higher PD tended to agree. Other research has examined differences in formulaic language used in persuasive essays to express opinions or address the reader. A recent study of persuasive essays from Korea, the UAE, and Western nations like the United States (Schenck, 2023) evaluated PD, providing further support for differences between writers from high and low PD countries. The study identified differences in how the reader was addressed by comparing stances, modals, and pronoun deixis. Stances, which are used to make expressions either more or less

forceful or assertive (Fatihi, 2019), can reveal the writer's authority and his or her attitude toward the reader. Modals also reflect authority and assertiveness, as in the following examples:

1. You (might/could) see a doctor.
2. You (should/ought to) see a doctor.
3. You must see a doctor. (Celce-Murcia & Larsen-Freeman, 1999, p. 146)

The degree of authority increases as modals move from examples one to three. In example three, the authors show the highest degree of assertiveness and authority. Finally, pronoun deixis reflects the distance between a writer and his or her reader. Using pronouns such as *I* and *you* denote a direct connection between the writer and the reader. In addition, the use of inclusive *we* denote a close connection (e.g., *We* need to end world hunger), suggesting that the writer is from the same speech community or cultural context as the reader (Petersoo, 2007; Yule, 1996). In contrast, relationships denoting "other" are suggested by pronouns like *he*, *she*, *it*, and *they* (Yule, 1996, p. 10). The distance expressed through these pronouns may reduce conflict, suggesting that the writer does not directly support a proposition.

Collective examination of stances, modals, and pronoun deixis may reveal insights about how ChatGPT deals with discourse related to PD in persuasive essays. Thus far, little research has been conducted to analyze ChatGPT's use of language related to PD. If limitations are identified, this information may be used by educators to design more effective pedagogical techniques that integrate AI-generated texts and feedback in classes designed to emphasize persuasive writing. Later, the research could be expanded to include other writing or speaking tasks, thereby improving how AI-generated text or feedback is used in language classrooms.

### **A Need for Further Examination of ChatGPT Writing**

The extent to which LLMs like ChatGPT can replicate culturally imbued language in persuasive texts is currently unknown. LLMs may utilize more generic forms of discourse to generate persuasive essays, lacking mechanisms needed to address a reader's contextual or cultural needs. A study of AI-generated text affirms this perspective, suggesting that ChatGPT does not have the pragmatic knowledge needed to adequately construct or assess writing (Algaraady & Mahyoob, 2023). Overall, deficiencies related to cultural aspects of written discourse reveal a need for further research on ChatGPT. By investigating how cultural values like PD are (or are not) reflected in writing, valuable insights about potential deficiencies of AI-generated persuasive texts and assessments may be gleaned. Educators can then critically assess ChatGPT output, designing instruction that heightens awareness of rhetorical variation, which is dependent upon unique cultural and situational contexts (Huang & Tan, 2023).

With a move to utilize ChatGPT as a tool for writing evaluation (Algaraady & Mahyoob, 2023; Fitria, 2023; Koraishi, 2023), investigation of AI-constructed rhetoric and discourse becomes even more essential. ChatGPT may rely on corpus data that promulgates culturally biased narratives. Concerning this issue, preliminary research suggests that ChatGPT expresses some alignment with American culture, yet lacks the cultural sensitivity needed to address learner issues from diverse cultural backgrounds (Cao et al., 2023).

The potential limitations of ChatGPT suggest that writing teachers must take a more active role in adapting AI technologies to classes that emphasize persuasive essays. However, inadequate examination of LLM essay construction limits understanding of learner needs, as well as associated pedagogical reforms required to address such needs. A comparison with authentic English texts is needed to provide a more complete perspective of how ChatGPT utilizes rhetoric associated with cultural values like PD. As pointed out by Crosthwaite and Baisa (2023), corpus data provides an effective means of identifying stylistic inconsistencies with register and genre within AI-generated essays. Concerning this issue, they state that:

*The authenticity of corpus data – that is, language data produced by humans – should be seen as a more reliable indicator of real language-in-use, particularly for second-language learners who do not have the benefit of being able to easily authenticate whether a given output would match what a native speaker of the target language might produce in the same context. (Crosthwaite & Baisa, 2023, p. 2)*

Comparison of ChatGPT with authentic corpus materials may reveal key differences in persuasive writing, which educators can use to evaluate and correct shortcomings associated with AI-generated text. Cultural preferences for language denoting PD may be revealed in English-L1 texts, yielding additional insights for enhancing pedagogical techniques that integrate AI. Information obtained from the present study may also help English teachers gain a heightened awareness of both AI writing systems and cultural norms. This awareness may then be imparted to learners, facilitating the acquisition of critical skills for the modification of written discourse that specifically addresses the reader.

### Research Questions

Due to a need to evaluate ChatGPT for potential inconsistencies related to language that denotes PD, the following questions were posed for the analysis of persuasive essays:

1. How do stances in ChatGPT essays differ from those found in essays written by L1-English writers? What may this difference reveal about ChatGPT's ability or inability to express varying degrees of power distance?
2. How do modals in ChatGPT essays differ from those found in essays written by L1-English writers? What may this difference reveal about ChatGPT's ability or inability to express varying degrees of power distance?
3. How does pronoun deixis in ChatGPT essays differ from that found in essays written by L1-English writers? What may this difference reveal about ChatGPT's ability or inability to express varying degrees of power distance?

Through investigating the questions above, it was hoped that valuable data could be obtained to facilitate the use of AI technologies in writing classrooms that emphasize persuasive writing.

### METHOD

An authentic English corpus was needed to compare ChatGPT's use of discourse related to PD. The International Corpus Network of Asian Learners of English (ICNALE) was an ideal choice for the present study. This 2-million-word corpus includes written samples of persuasive essays from L1-English learners, as well as a range of English as a Foreign Language (EFL) learners from countries throughout Asia (Ishikawa, 2018). It also includes clear guidelines for data collection, which could be used to direct ChatGPT as it was tasked with replicating the corpus. Finally, highly standardized data collection ensures that essays represent a similar genre and length, thereby increasing the validity of comparison.

All participants who contributed to ICNALE were asked to write essays using two prompts (Ishikawa, 2013, p. 97):

Topic 1: It is important for college students to have a part-time job.

Topic 2: Smoking should be completely banned at all restaurants in the country.

After receiving a prompt, each learner was given 20 to 40 minutes to write an essay from 200 to 300 words. No dictionaries or other reference tools were allowed (Ishikawa, 2013). A specific

evaluation of a persuasive writing task could be conducted by controlling the topic in this way. A total of 200 essays written by 100 L1-English university learners were utilized for analysis (100 essays about a part-time job and 100 essays about smoking). Learners ranged in age from 19 to 29 and studied various subjects in the humanities and sciences. They came from several different English-speaking countries, including the United States, Great Britain, Canada, Australia, and New Zealand. These countries have low scores for PD, which was considered as the essays were evaluated (Northouse, 2021).

To create a corpus for comparison, ChatGPT-3.5 was asked to write 100 different essays for Topics 1 and 2 (the topics used to construct the ICNALE corpus) for a total of 200 essays. Since ChatGPT would not write so many essays at one time, the LLM tool was asked to write five different essays at a time until the desired number of essays was reached. The following two prompts were used to obtain the essays in the corpus.

1. Write five different essays from 200 to 300 words about the following topic: It is important for college students to have a part-time job.
2. Write five different essays from 200 to 300 words about the following topic: Smoking should be completely banned at all the restaurants in the country.

After the two corpora were prepared, types and tokens were compiled into Table 1.

As revealed by a similar number of tokens, the two corpora were equivalent in size, making the comparison of formulaic language and vocabulary possible. At the same time, ChatGPT had 37% fewer types, which suggests that word variation was more limited than ICNALE.

### Indicators of Power Distance

Power Distance (PD) represents the inequality of the relationship between the writer and his or her reader. This inequality can be expressed by heightening the writer's authority (e.g., using authoritative language about assertions) or excluding the reader from discourse (e.g., failing to refer to the reader through inclusive language). To examine PD, stances, modals, and pronouns were collectively examined.

Following the preparation of corpora, words that could denote differences in power distance were evaluated. First, stances were examined. To obtain a general idea of how ChatGPT uses stances to assert authority, three exemplars given in a study by Min et al. (2019) were used for each of the six following categories: epistemic certainty adjectives, epistemic likelihood adjectives, epistemic certainty adverbs, epistemic likelihood adverbs, epistemic certainty verbs, and epistemic likelihood verbs (Table 2).

By no means do the selected stances represent a complete list. However, as exemplars, they represent commonly used stances that may yield insights as to how ChatGPT formulates discourse.

Next, modals were chosen for analysis of a writer's authority. A total of ten modals were examined using the following categories of authority proposed by Celce-Murcia and Larsen-Freeman (1999, p. 146): highest authority (*must, have to, has to, got to, need to*), intermediate authority (*should, had better, ought to*) and lowest authority (*could, might*).

Table 1. Size of corpora used for study

	ICNALE Corpus	ChatGPT Corpus
Types	3356	2108
Tokens	45028	43782

Table 2. Epistemic stances chosen for examination

Epistemic Certainty Adjectives	Epistemic Likelihood Adjectives	Epistemic Certainty Adverbs
Apparent	Likely	Actually
Certain	Possible	Certainly
Obvious	Probable	Definitely
Epistemic Likelihood Adverbs	Epistemic Certainty Verbs	Epistemic Likelihood Verbs
Apparently	Conclude	Assume
Perhaps	Notice	Guess
Probably	Prove	Seem

Finally, pronoun deixis was analyzed to give an idea of the distance maintained by the writer concerning his or her reader. The pronouns *I, you, we, he, she,* and *they* were examined along with associated demonstrative pronouns *my, your, our, his, her,* and *they*. Pronouns like *I, you,* and *we* represent a close relationship between the writer and their reader. In contrast, pronouns such as *he, she,* and *they* represent general examples or something “other,” which denotes a degree of distance between the writer and their reader (Yule, 1996). Words like *woman, man, women, men, person,* and *people* may also describe general examples, establishing a more distant relationship between the writer and reader. Therefore, these words were included in the examination.

### Evaluation of Corpus Data

After corpora were compiled, indicators of PD were tallied from each corpus. The AntConc concordance tool was used to search for frequencies of words or phrases from each corpus. After entering values into a spreadsheet, data from the ChatGPT corpus were statistically compared to data from ICNALE for each group of PD indicators (stances, modals, and pronouns). The Wilcoxon signed ranks test for nonparametric data was used because the number of words for each category varied from 10 to 18, a small number where normal distributions could not be assumed. Following the Wilcoxon signed ranks test calculation, frequencies for each category were graphically charted and evaluated. Empirical values were also evaluated through a qualitative view of the essays, thereby lending new insights into how ChatGPT uses different forms of discourse related to PD.

## RESULTS AND DISCUSSION

### Epistemic Stances

Comparison of epistemic stances revealed significant differences according to the Wilcoxon signed ranks test ( $z = -3.411$ ;  $p = .001$ ). The difference in stances was also clearly evident when the data was examined qualitatively (See Table 3 and Appendix A).

As revealed in Table 3, common stances were rarely used by ChatGPT. Only the stance *likely* was used twice. Other stances such as *possible, actually, prove,* and *seem* were used only once. These stances were used in the following ways in ChatGPT:

1. Students should prioritize their academic commitments and seek flexible work arrangements when **possible**.
2. Part-time jobs can **actually** have a positive impact on their studies.
3. These connections can **prove** instrumental in providing references, job recommendations, and even potential career opportunities in the future.

Table 3. Frequencies of epistemic stances in the ICNALE and ChatGPT corpora

Epistemic Stance	ICNALE Corpus	ChatGPT
1. Apparent	0	0
2. Certain	5	0
3. Obvious	10	0
4. Likely	5	2
5. Possible	13	1
6. Probable	1	0
7. Actually	14	1
8. Certainly	4	0
9. Definitely	13	0
10. Apparently	2	0
11. Perhaps	8	0
12. Probably	29	0
13. Conclude	0	0
14. Notice	2	0
15. Prove	1	1
16. Assume	2	0
17. Guess	6	0
18. Seem	20	1

4. While balancing work and studies may **seem** daunting, it is essential for students to establish a proper routine and set realistic goals.

Examples reveal generic means to address the topic of part-time jobs. Furthermore, each stance expresses epistemic likelihood, which lessens a writer’s authority over an assertion. The strength of stances was further weakened by the modals *can* or *may*, as in examples 2 through 4. Whereas the generic quality of discourse can preclude close, personal connections with a reader (heightening PD), weak stances may limit an author’s authority (lowering PD).

Stances from the English corpus were varied in style and strength of epistemic certainty. More stances of epistemic certainty were utilized. The following are some examples:

1. That would be the democratic thing to do in my view and I am **certain** that most people would have a lot more respect for the government rather than being told to do this and do that all the time.
2. For these reasons and many more, I think it is **obvious** that college students should not have to have a part-time job.
3. With this kind of **obvious** link between a life-threatening disease and a stupid habit, the choice to quit smoking and the choice to ban smoking at restaurants would seem to be a good one.
4. For other students with technical majors, I think the case is similar, and they should find some way to incorporate part-time work into their curriculum because it will improve their prospects of employment when they go about seeking a full-time position.



Epistemic stances from the L1-English corpus were often stronger. As in the case of Example 1, epistemic certainty is used to exert authority in conjunction with the pronoun *I*, which establishes a connection with the reader. Stances are also used with other auxiliary verbs to mitigate the strength of epistemic certainty, as in the case of Example 2 (“I think it is obvious”). While both ChatGPT and ICNALE reveal a lessening of PD by adding auxiliary verbs or phrases, only L1-English writers can utilize stances at varying degrees of intensity. In addition, L1-English learners use stances to address potential opponents of an idea, as in the case of “While I **guess** that there might be a huge backlash at first, people would adopt the idea soon enough and a ban can only help the health situation in the long run.” Overall, sophistication in the essays of L1-English learners suggests that ChatGPT cannot modify the intensity of rhetoric. The reader is only addressed through generic stances of epistemic likelihood, which make an argument appear dull and unimportant. Whereas L1-English writers demonstrate a complex awareness of PD, ChatGPT utilizes standardized collocations of a more general nature, which ultimately fail to convince or connect.

### Modals of Authority

Research Question Two, which sought to examine differences in the force of modals, also revealed significantly different results (See Appendix B). The Wilcoxon signed ranks test revealed a Z value of -2.100, which was significant at  $p = .036$ . As with stances, the word variety of modals was limited. Table 4 reveals more information.

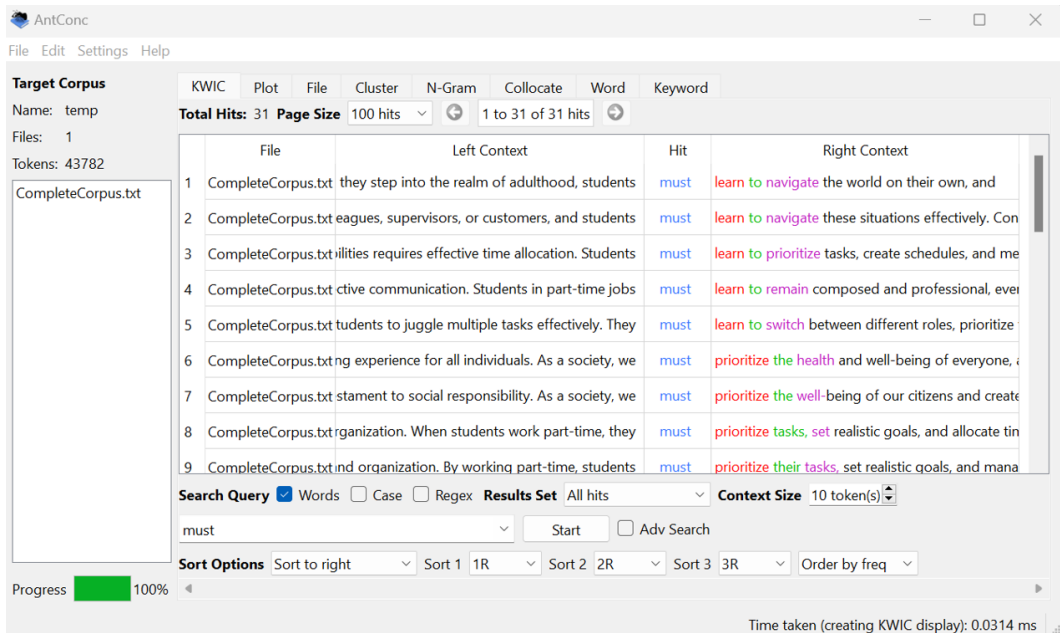
While ChatGPT used weak stances that avoided epistemic certainty, it also used more forceful modals like *must*, *have to*, and *need to*. The word *must* be used extensively, suggesting the writer was taking a more forceful stance concerning PD. However, usage of the modal was highly standardized, as revealed by the most common collocations (Figure 1).

The same words and collocations were recycled, referring to common topics related to either students or society. Although authoritative modals were used, they were associated with non-contentious issues (e.g., the importance of health). ChatGPT chose topics that the reader would obviously agree with. Whereas generic terms seem to increase the distance from a reader, which is a sign of higher PD, avoidance of contention appears to reflect a lower degree of authority, which may reflect lower PD. Concerning *should*, which was used 113 times by ChatGPT, recycling, and collocations were even more standardized. Phrases like “should not compromise”, “should not negatively impact”, and “should not actively seek” were extensively utilized. Overall, high and intermediate authority modals were used with standardized arguments that could not easily be debated.

Table 4. Frequencies of modals in the ICNALE and ChatGPT corpora

Modal	ICNALE Corpus	ChatGPT
1. could	35	1
2. got to	1	0
3. had better	0	0
4. has to	5	0
5. have to	89	1
6. might	31	3
7. must	14	31
8. need to	31	1
9. ought to	0	0
10. should	268	113

Figure 1. Examples of the “must” modal in the ChatGPT corpus



Because ChatGPT must rely on corpus data to select the “right” modal, standardized collocations in common situational contexts appear overutilized. This results in the construction of generic arguments about concepts such as tasks, goals, and health, which constitute ideas that are difficult to refute. For example, no one may argue with the idea that students “must prioritize tasks.” This reflects a lack of interaction with the argument and the reader. Albeit limited, there were attempts to directly interact with the reader through statements like “We must prioritize the well-being of our citizens.” Although inclusive “We” is utilized, the statement is not controversial, ensuring that the reader will agree with the accuracy of the modal.

As revealed in Table 4, modals utilized by L1-English writers were much more varied, reflecting unique contexts and opinions. Collocations associated with modals were much more sophisticated, revealing a lot of hypothetical situations to support an argument (e.g., “If young people must have curricular activities...”/“You might be able to...”). They also used high authority modals with another language to lessen authority, thereby maintaining low PD, a characteristic of Western discourse. In the example, “I hope that one day soon we will not have to worry about it anymore,” an authoritative modal has to is transformed into a hypothetical situation by adding the modal will. Such hypothetical situations appear to reduce the force of an authoritative modal, thereby lowering PD. Expressions like I hope and we also lessen distance by cultivating a personal connection and common interest. Collectively, modals and associated collocations of L1 English learners revealed a much more cogent ability to strengthen assertions while simultaneously maintaining culturally appropriate levels of PD. This ability appeared lacking in essays written by ChatGPT.

## Pronoun Deixis

Concerning research question three, which sought to examine pronoun deixis, the Wilcoxon signed ranks test yielded a value of -1.917, which was nearly significant at the .05 level ( $p = .055$ ).

Like stances and modals, pronouns lacked variety in the ChatGPT corpus (See Table 5 and Appendix C). No examples of *I* or *you* were utilized to connect with the reader more directly. This reflects distance or high PD about the reader. The pronoun *we* is used extensively, suggesting an

Table 5. Frequencies of pronouns and common nouns in the ICNALE and ChatGPT corpora

Pronoun	ICNALE Corpus	ChatGPT
1. I	1015	0
2. You	297	0
3. He	24	0
4. She	12	0
5. They	466	103
6. We	155	271
7. My	247	0
8. Your	96	0
9. His	19	0
10. Her	7	0
11. Their	340	594
12. Our	70	107
13. women	0	3
14. men	0	0
15. woman	2	0
16. man	5	0
17. people	279	53
18. person	13	0

attempt to interact with the reader, thereby cultivating a closer relationship (low PD). However, as in prior examples of discourse used by ChatGPT, responses using this pronoun are highly formulaic, which may represent the training corpus from which it was derived. The collocations “we demonstrate our commitment” and “we create a more” were exclusively used with concepts such as society (e.g., an inclusive society) or health. ChatGPT appears to choose types of discourse based on consistency in the corpus used for AI training. Therefore, while writing seems impressive superficially, it lacks variety and unique opinions. It cannot also closely connect with the reader on a personal level. A lack of nouns concerning gender further illustrates this perspective. Only the word *women* was used three times in reference to the adverse impact of smoking on pregnancy. Except for these three examples, the complexities related to gender about getting a part-time job or aspects of smoking were disregarded.

Regarding the L1-English learner corpus, pronouns revealed great complexity in their use. The personal pronoun *I* was used to give personal opinions (e.g., I think the government should do more to regulate this area.), to present a novel personal perspective (e.g., So, now that I have a new perspective...), to present personal opinions (e.g., I have no sympathy for the people who choose to destroy their bodies by smoking cigarettes), and describe personal experiences (e.g., I have an easy schedule this semester, so....). Such a discussion makes the argument personal and conveys a unique perspective, thereby revealing ownership and support for the argument being made. The pronoun *you* was also used extensively to involve the reader, showing potential positive outcomes of a chosen argument (e.g., You can quickly adapt to new situations. / You can stay focused on your studies.). Pronouns like *he* or *she* were used to explain unique scenarios to support the point, as in the following example:

*What of the children? It's bad enough that Granny's emphysema attacks worsen with every Sunday brunch, as clouds of noxious fumes from nearby tables create a deadly nimbus of toxic pollution above those hennaed waves of hair. What of the infant, his life irremediably shortened with each contaminated breath he takes?*

The above excerpt explains hypothetical figures from different genders and ages, connecting with a wide readership. Overall, the use of pronouns by L1-English writers was much more complex, reflecting a clear attempt to connect with the reader. A direct connection appeared to lessen social distance or formality, reflecting low PD. Pronouns were carefully selected to cultivate a personal relationship, helping the reader identify with the proposed arguments.

## CONCLUSION

Results suggest that ChatGPT lacks an awareness of power distance, which impacts essay construction. Some elements of high PD (e.g., authoritative modals) are used, yet they reflect a standardized argument that precludes disagreement, thereby lessening PD. Concerning pronoun deixis, extensive utilization of the all-inclusive *we* also reflect a generic approach designed to lessen PD. Results reduce an author's authority, thereby lessening asymmetrical relationships between the writer and reader. Rather than representing a sophisticated mechanism for constructing discourse, findings appear to reflect cultural biases imbued in training input, which originated from low PD Western contexts, such as the United States.

In contrast to the ChatGPT corpus, compositions by L1-English learners from ICNALE vary considerably in the use of stances, modals, and pronouns. Such variation reveals a more complex relationship with the reader, where more contentious assertions are carefully supported. As an example, personal opinions and experiences are used to connect closely with the reader while bolstering a proposition, thereby lowering PD. This finding has distinct implications for writing pedagogy. Teachers must carefully select and emphasize authentic examples of rhetoric (e.g., from corpus data), which outline a spectrum of assertions and connect with readers via combinations of stances, modals, and pronouns. In addition, teachers will need to help learners address the cultural complexities of discourse in conjunction with additional concerns like gender or age. Finally, teachers will need to help learners avoid generic arguments characteristic of AI-generated persuasive texts, allowing the learner to take ownership of arguments and assert opinions by utilizing personal experience or novel perspectives. If supplemental AI technologies are provided with this kind of support, the potential of ChatGPT may be realized. As the results of the present study suggest, English teachers will need to carefully monitor and enhance AI-generated output, helping learners understand the complexities of cultural discourse within human compositions.

Results reveal that AI-generated texts have distinct limitations regarding culturally imbued language in persuasive essays. This finding has important implications for both first and second-language English classrooms. Learners in these classrooms must be provided with supplemental materials or explicit instruction to overcome the shortcomings of AI-generated texts. While the present study examined only one specific task, results appear to have implications for many writing tasks and cultural contexts. Additional study will be needed to develop a more comprehensive perspective of how AI-generated writing utilizes culturally imbued language. Such research may help educators understand the extent to which authentic English texts and explicit instruction are needed in language classes that integrate AI technologies.

## Limitations and the Need for Further Study

The results of the present study provided insights concerning how ChatGPT utilizes culturally imbued language. Although these insights may improve writing pedagogy, key limitations must be considered. The study examined a persuasive task of very limited scope. In future research, various

writing tasks from different pragmatic contexts (e.g., legal documents and cover letters) should be examined, providing a more comprehensive perspective on how ChatGPT addresses the reader.

Secondly, the present study examined only L1-English learners' usage of meta-discourse related to PD. The information obtained from the analysis can also provide EFL teachers with valuable insights into how discourse can be constructed to communicate with readers from L1-English contexts. At the same time, it does not address readers from other cultural or linguistic backgrounds. Aspects of language related to PD may vary in contexts like South Korea and the UAE, where power distance is highly valued. Additional research is needed to help educators understand how learners can use AI to write cogent, convincing, and culturally appropriate texts in diverse contexts.

Although research has yielded some key insights, additional study of meta-discourse is needed so that PD and other cultural aspects of writing composition can be better understood. A more comprehensive examination may provide additional insights concerning the enhancement of AI-instruction or evaluation.

## REFERENCES

- Algaraady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journal Special Issue on CALL*, 9(9), 3–17. doi:10.24093/awej/call9.1
- Biber, D. (1996). *University language: A corpus-based study of spoken and written registers*. John Benjamins Publishing.
- Bjørge, A. K. (2007). Power distance in English lingua franca email communication. *International Journal of Applied Linguistics*, 17(1), 60–80. doi:10.1111/j.1473-4192.2007.00133.x
- Bogost, I. (2022, December 16). ChatGPT is dumber than you think. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/>
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL / EFL teacher's course* (2nd ed.). Heinle & Heinle Publishers.
- Crosthwaite, P., & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast! *Applied Corpus Linguistics*, 3(3), 1–4. doi:10.1016/j.acorp.2023.100066
- Fatihi, A. R. (2019). *Waggish coquetry in South Asian street communication*. Lexington Books.
- Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching*, 12(1), 44-58. doi:10.15294/elt.v12i1.64069
- Halliday, M. A. K., & Hasan, R. (1989). *Language, context and text: Aspects of language in a social semiotic perspective*. Oxford University Press.
- Hamadouche, M. (2013). Intercultural studies in the Arab world from a contrastive rhetoric perspective. *Arab World English Journal*, 2, 181–188.
- Hamam, D. (2020). A study of the rhetorical features and the argument structure of EAP essays by L1 and L2 students in the UAE. *Journal of Asia TEFL*, 17(2), 699–706. doi:10.18823/asiatefl.2020.17.2.28.699
- Hammad, H. M. (2002). *A contrastive examination of the rhetorical patterns of Arab speakers' Arabic and ESL writing* [Master's thesis, University of Calgary]. PRISM. <http://hdl.handle.net/1880/42480>
- Huang, J., & Tan, M. (2023). The role of ChatGPT in scientific communication: Writing better scientific review articles. *American Journal of Cancer Research*, 13(4), 1148–1154. PMID:37168339
- Hyland, K. (2018). *Metadiscourse: Exploring interaction in writing*. Bloomsbury Publishing.
- Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In S. Ishikawa (Ed.), *Learner corpus studies in Asia and the world* (pp. 91–118). Kobe University Press. [http://www.lib.kobe-u.ac.jp/infolib/meta\\_pub/G0000003kernel\\_81006678](http://www.lib.kobe-u.ac.jp/infolib/meta_pub/G0000003kernel_81006678)
- Ishikawa, S. (2018). *ICNALE: The international Corpus network of Asian learners of English*. Kobe University Press. <http://language.sakura.ne.jp/icnale>
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873–15892. doi:10.1007/s10639-023-11834-1
- Jiang, X. (2006). Cross-cultural pragmatic differences in US and Chinese press conferences: The case of the North Korea nuclear crisis. *Discourse & Society*, 17(2), 237–257. doi:10.1177/0957926506060249
- Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, 3(1), 55–72.
- Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1–14. doi:10.1080/17501229.2023.2240316

- Liu, Y., & Du, Q. (2018). Intercultural rhetoric through a learner lens: American students' perceptions of evidence use in Chinese yìlùn wén writing. *Journal of Second Language Writing, 40*, 1–11. doi:10.1016/j.jslw.2018.01.001
- Min, S., Paek, J. K., & Kang, Y. (2019). Exploring the use of hedges and stance devices in relation to Korean EFL learners' argumentative writing Qualities. *English Teaching, 74*(1), 3–23. doi:10.15858/engtea.74.1.201903.3
- MLA-CCCC Joint Task Force on Writing and AI. (2023). *MLA-CCCC Joint Task Force on Writing and AI working paper: Overview of the issues, statement of principles, and recommendations*. Modern Language Association. <https://commons.org/app/uploads/sites/1003160/2023/07/MLA-CCCC-Joint-Task-Force-on-Writing-and-AI-Working-Paper-1.pdf>
- Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: Perceptions of EFL Faculty Members. *Education and Information Technologies, 1*–23. doi:10.1007/s10639-023-11917-z
- Northouse, P. G. (2021). *Leadership: Theory and practice*. Sage Publications.
- Orts, M. Á. (2016). Power distance and persuasion: The tension between imposition and legitimation in international legal genres. *Journal of Pragmatics, 92*, 1–16. doi:10.1016/j.pragma.2015.11.009
- Paltridge, B. (2012). *Discourse analysis: An introduction* (2nd ed.). Bloomsbury. doi:10.5040/9781350934290
- Pérez-Paredes, P. (2023). *Using ChatGPT (twice) to answer one of my own undergrad essay titles*. University of Murcia. Retrieved from <https://www.um.es/languagecorpora/2023/02/09/using-chatgpt-to-answer-one-my-own-undergrad-essay-titles/>
- Petersoo, P. (2007). What does 'we' mean? National deixis in the media. *Journal of Language and Politics, 6*(3), 419–436. doi:10.1075/jlp.6.3.08pet
- Schenck, A. (2023). Examining the impact of subtle cultural differences on rhetorical variation in EFL essays. *Acuity: Journal of English Language Pedagogy, Literature and Culture, 8*(1), 1-16. <https://files.eric.ed.gov/fulltext/EJ1374876.pdf>
- Shah, R., Gao, Z., & Mittal, H. (2014). *Innovation, entrepreneurship, and the economy in the US, China, and India: Historical perspectives and future trends*. Academic Press.
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*(11), 13943–13967. doi:10.1007/s10639-023-11742-4
- Yule, G. (1996). *Pragmatics*. Oxford University Press.

## APPENDIX A

### Statistical evaluation of stances

		N	Mean Rank	Sum of Ranks
English - ChatGPT	Negative Ranks	0 <sup>a</sup>	.00	.00
	Positive Ranks	15 <sup>b</sup>	8.00	120.00
	Ties	3 <sup>c</sup>		
	Total	18		
a. English < ChatGPT				
b. English > ChatGPT				
c. English = ChatGPT				

		English - ChatGPT
Z		-3.411 <sup>b</sup>
Asymp. Sig. (2-tailed)		.001
a. Wilcoxon Signed Ranks Test		
b. Based on negative ranks.		

## APPENDIX B

### Statistical evaluation of modals

		N	Mean Rank	Sum of Ranks
English - ChatGPT	Negative Ranks	1 <sup>a</sup>	3.00	3.00
	Positive Ranks	7 <sup>b</sup>	4.71	33.00
	Ties	2 <sup>c</sup>		
	Total	10		
a. English < ChatGPT				
b. English > ChatGPT				
c. English = ChatGPT				

		English - ChatGPT
Z		-2.100 <sup>b</sup>
Asymp. Sig. (2-tailed)		.036
a. Wilcoxon Signed Ranks Test		
b. Based on negative ranks.		



## APPENDIX C

### Statistical evaluation of pronouns

		N	Mean Rank	Sum of Ranks
English - ChatGPT	Negative Ranks	4a	9.00	36.00
	Positive Ranks	13b	9.00	117.00
	Ties	1c		
	Total	18		
a. English < ChatGPT				
b. English > ChatGPT				
c. English = ChatGPT				

	English - ChatGPT
Z	-1.917 <sup>b</sup>
Asymp. Sig. (2-tailed)	.055
a. Wilcoxon Signed Ranks Test	
b. Based on negative ranks.	

*Andrew Schenck has taught English for over 20 years. While working, he has managed to earn a master's degree in TESOL, a master's degree in Educational Foundations, an Ed.D. in Educational Leadership with a specialization in English, and a Ph.D. in English Language and Literature. Throughout his career, he has focused on publishing research concerning English literature, explicit grammar instruction, and language acquisition.*