

# NATIVE, NONNATIVE, OR BILINGUAL? A CONCISE ASSESSMENT OF CHATGPT'S SUITABILITY FOR SECOND-LANGUAGE INSTRUCTION AS A NATIVE OR NONNATIVE PEDAGOGUE

Noboru Sakai, Tamagawa University, Tokyo Keizai University, St. Marianna University School of Medicine, Tokyo University of Technology

---

## ABSTRACT

*This study aims to investigate ChatGPT's ability to comprehend input from nonnative speakers, specifically those learning English as a second language, with Japanese speakers serving as the model population. The experiment examines how ChatGPT evaluates the difficulty levels of the Test of English for International Communication (TOEIC), which is widely taken by nonnative English speakers of varying proficiency levels. This study also measures ChatGPT's response to discourses produced by nonnative speakers—one containing grammatical errors and the other incorporating sociolinguistic or strategic competence expressions. The findings indicate that ChatGPT demonstrates proficiency in lexical and grammatical features and can comprehend nonperfect grammatical discourses produced by nonnative speakers. However, ChatGPT does not accurately perceive the varying degrees of difficulty experienced by nonnative speakers. Moreover, while ChatGPT can comprehend sociolinguistic expressions when the context is clear, its understanding of other communication strategies employed by nonnative speakers is relatively limited.*

**Keywords:** ChatGPT, native vs. nonnative, linguistic tolerance, Computer Assisted Language Learning, individual study

## INTRODUCTION

The burgeoning prominence of ChatGPT (by OpenAI: <https://openai.com/blog/chatgpt>), following its public unveiling in November 2022, has engendered significant interest among both learners and teachers in the realm of language acquisition. Its versatile applications and capacity for facilitating diverse educational endeavors have been met with considerable approbation. However, this enthusiasm is tempered by concerns regarding its potential deleterious impact on students' cognitive development, as ChatGPT's ability to provide cogent responses with markedly reduced effort

when juxtaposed with traditional computer-assisted language activities, such as machine translation, may inadvertently impede their learning progress.

Numerous educational establishments have expeditiously adopted policies to ban the utilization of ChatGPT within their respective institutions. For instance, Sciences Po, a preeminent academic institution in France, established a tripartite framework for ChatGPT implementation predicated upon an overarching proscription without particular stipulations (2023). This predilection appears to be increasingly pervasive in adherence to principles of academic integrity, particularly within language

courses, where the caliber of output is conspicuously characterized by a standardized variation.

Potential apprehensions exist, yet it is highly probable that an inundation of recommendations, as well as both favorable and adverse accounts regarding ChatGPT's application in education, will transpire. Consequently, the realm of language acquisition, specifically computer-assisted language learning, will undoubtedly become saturated with research investigations predominantly influenced by ChatGPT. Thus, the efficacious utilization of ChatGPT emerges as a paramount subject of inquiry.

In the pursuit of this intellectual enterprise, a pivotal topic of investigation concerns the function of ChatGPT as a linguistic exemplar, specifically examining whether it operates as (a) an instructive agent analogous to a native language speaker imparting knowledge to nonnative speakers or (b) a pedagogue of nonnative speakers who have attained exceptional proficiency in the target language.

As a linguistic model, ChatGPT abstains from engaging in idiosyncratic variations exhibited by nonnative speakers, as such interactions could potentially ostracize these individuals when juxtaposed with their native-speaking counterparts (Sakai, 2023). Consequently, the linguistic underpinnings of ChatGPT appear predominantly anchored in native speaker norms, and the vantage points of nonnative language learners remain largely unexplored within the prevailing paradigm. Thus, this research seeks to investigate the efficacy of ChatGPT in fostering language acquisition among nonnative speakers, with a particular emphasis on English as an exemplar, given its pre-eminent status within the ChatGPT framework.

## LITERATURE REVIEW

Undeniably, ChatGPT has garnered significant attention within the domain of Computer-Assisted Language Learning (CALL), as evidenced by the emergence of preliminary academic published investigations. Ali et al. (2023) presented an avant-garde investigation that posits potential advantages of incorporating ChatGPT as a linguistic instrument for heterogeneous cohorts, spanning from elementary school students to professional training attendees in both individual and collective educational environments. Simultaneously, they proffer

a prudential warning concerning the unwarranted dependency on this linguistic paradigm, urging circumspection from both learners and educators alike to avoid overreliance on it.

The extant research exhibits an encouraging ambiance towards ChatGPT, potentially attributable to widespread anticipation. Concurrently, it is imperative to contemplate potential drawbacks. For instance, Vera and Laurence (2022) comprehensively delineated the detriments associated with machine translation, which are likely to parallel the experience in second language (L2) acquisition using ChatGPT. These disadvantages encompass students' exploitation of the system for their scholastic tasks, apprehensions regarding machine translation's potential to overshadow traditional language pedagogy and learning in terms of proficiency augmentation, disquietudes about the insufficiency of nurturing cognitive aptitudes through machine translation utilization, and trepidations of subpar outcomes resulting from dependence on machine translation. All these aspects may not be fully applicable, but it is imperative to acknowledge the potential disadvantages that warrant consideration, as they encompass more sophisticated and intricate aspects of machine translation and ChatGPT in particular.

To delve more profoundly into the subject matter, this scholarly composition commences with an exploration of Computer-Assisted Language Learning (CALL) utilization in the context of test preparation, as it constitutes a primary application of CALL. Specifically, this section elaborates upon aspects pertinent to the Test of English for International Communication (TOEIC), which is extensively administered to nonnative speakers, even those without aspirations to reside in English-speaking nations. Consequently, it is anticipated that the research and pedagogical practices encompass a vast spectrum of language proficiency levels. Furthermore, the aggregation of data from nonnative speakers is accessible, which proves invaluable in investigating the potential employment of ChatGPT as an English instructor for this demographic.

Kim (2022) investigated the utilization of artificial intelligence (AI) within the context of TOEIC classes. This study uses an AI system named Soljam that evaluates learners' proficiency through answering several questions and further provides

questions that are appropriate for the students' proficiency level that has been assessed. The study concludes that the incorporation of AI technology markedly impacts TOEIC proficiency. This research serves as a cornerstone when evaluating the potential applicability and efficacy of ChatGPT in the realm of language acquisition.

More broadly, a crucial aspect of CALL in the domain of language acquisition lies in its capacity to foster motivation, enabling learners to engage in self-directed study and facilitating a personalized educational experience. In this regard, CALL constitutes a potent instrument for augmenting language acquisition throughout various phases of language study, encompassing initiation, progression, and culmination (Sakai, 2007). Within the realm of TOEIC preparation, Wijayanti et al. (2019) discovered that 69% of participants exhibited a favorable disposition towards online learning. Nevertheless, efficacious implementation of CALL for language acquisition is not devoid of obstacles, as Japanese students grapple with technology assimilation and self-assurance in employing it for language study (Ozawa, 2019). To surmount these hurdles, Cheng (2020) proffered a web-based, game-oriented methodology that considerably enhanced students' motivation and self-confidence, as it allows students to improve their proficiency through an enjoyable learning environment, enabling them to study more effectively than through other learning methods.

In order to surmount apprehension and immerse oneself in protracted intellectual development, a pivotal element influencing the efficacy of ChatGPT as a linguistic support apparatus is the learners' perception of its role in language acquisition. Within Second Language Acquisition (SLA), the discourse concerning the efficacy of native-speaker teachers versus nonnative-speaker teachers has endured for an extended period, and a consensus among scholars remains elusive (Selvi, 2011). Despite this, the prevailing tendency in the field is to acknowledge and respect the merits and limitations of both teacher categories (Matsuda & Matsuda, 2001). Native teachers continue to command significant precedence in oral skills, which is also acknowledged by nonnative teachers (Árva & Medgyes, 2000). Clark and Paran (2007) revealed that, in addition to teaching experiences and educational background, native-speaker

proficiency is considered a significant criterion in the hiring process.

Native Speaker Teachers (NST) may proffer benefits such as ameliorating oral communication skills, enriching vocabulary, and addressing cultural nuances (Mahboob, 2004). However, they may not consistently furnish comprehensive responses to students' inquiries. In contrast, Non-Native Speaker Teachers (NNST) demonstrate superior aptitude in grammar, pedagogical methodologies, and offering emotional support to students, given their own experiences as language learners (Mahboob, 2004). Calafato (2019) posited that, in addition to traditional perspectives, NNSTs have garnered recognition as being multilingually proficient, thus highlighting their unique strengths.

The students' perception regarding their instructor constitutes a crucial element in determining the efficacy of language acquisition. Lasagabaster and Sierra (2002) examined students' predilections for NST and NNST, with the findings suggesting that students generally favored NST or a combination of NST and NNST. However, within higher education, NSTs were marginally preferred due to their assessment capabilities. In a separate study, Todd and Pojanapunya (2009) reported that students explicitly favored NSTs, but implicit preferences did not exhibit a significant distinction between NST and NNST. Floris and Renandya (2020) underscored the importance of prioritizing teachers' competencies, qualifications, and experiences, rather than solely emphasizing the NST versus NNST dichotomy when selecting language instructors.

In the contemporary milieu of English as an international language, encompassing World Englishes and English as a Lingua Franca, the perspectives on language instructors, be they native or nonnative speakers, have evolved. Rather than solely serving as arbiters of correctness, these instructors now embody a more nuanced role, signifying the "authenticity" of linguistic productions (Kemaloglu-Er & Lowe, 2022). Consequently, responses to language teachers have undergone significant transformation. The position of Artificial Intelligence within this changing landscape presents a compelling subject for further inquiry.

Whether the ChatGPT model exhibits characteristics akin to native speakers (NST) or nonnative speakers (NNST) of English, a salient attribute of

the system is its adeptness in managing nonnative or erroneous variations of language. In this context, the trajectory of technological advancements appears promising. For instance, Matassoni et al. (2018) documented their integration of nonnative children into their deep neural network, culminating in the enhancement of the target language's specificity. The proliferation of technology is experiencing an expeditious ascent in contemporary times, consequently leading to enhanced system performance. In this context, ChatGPT's adeptness in addressing the matter at hand demonstrates a commendable level of proficiency.

Moreover, to effectively comprehend discourse from nonnative speakers, it is crucial to possess communicative competence that extends beyond mere understanding of grammatical structures or lexical meanings. This is because nonnative speakers often employ various techniques to convey their intended message. In this context, the notion of communicative competence, initially posited by Hymes (1972), encompasses a comprehensive communication aptitude that includes both sociolinguistic and strategic competences. This concept has evolved and been adapted over time, emphasizing that effective communication involves more than just grammatical accuracy. Notwithstanding, there appears to be a scarcity of research in this domain, and this study aims to contribute to its advancement.

Concerning the sociolinguistic factor, one significant challenge lies in deciphering nongrammatical or extra-textual expressions. Numerous scholars, such as Miyake (2007) and Sakai (2013), have demonstrated that the utilization of such extra-linguistic elements is governed by specific underlying mechanisms, even when considering solely Japanese and English. In the context of multilingual communication involving nonnative input, Sakai (2015) posited that it is exceedingly arduous to accurately translate the meaning of such expressions from one language (A) to equivalent expressions in another language (B) while preserving their intuitive essence. Consequently, evaluating the extent to which ChatGPT can address sociolinguistic aspects presents an enduring inquiry.

Based on the discussion above, the investigation of the present study delves into the efficacy of ChatGPT in relation to input that mirrors the linguistic intuitions of nonnative English speakers, with an emphasis on Japanese learners, who have been identified as members of a cohort that faces formidable challenges in acquiring English proficiency.

## RESEARCH QUESTIONS AND METHODOLOGY

This scholarly probe scrutinizes the outputs engendered by ChatGPT versions 3.5 and a marginally augmented iteration of 4.0 when interacting with Japanese English-language learners. The research questions underpinning this study are delineated as follows:

1. In what manner does ChatGPT evaluate multifarious facets of linguistic intricacy, specifically with respect to discerning gradations of difficulty grounded in the perceptions of Japanese speakers?
2. To what degree is ChatGPT capable of accommodating and apprehending input proffered by Japanese speakers?

To address the initial research question, which pertains to the identification of linguistic arduousness, particularly from the vantage point of nonnative speakers, this research further defines the actual research task with the question: Can ChatGPT accurately assess what nonnative speakers consider a difficult question, thereby ensuring synchronization with the test takers' perceptions of difficulty? To answer the question, this analysis employs an array of TOEIC questions, which has undergone copious investigation concerning the characteristics demarcating each level of difficulty. Moreover, a plethora of Japanese pedagogical materials is predicated upon these score stratifications. This study undertakes an analysis of two specific volumes, selected from among a range of literary works:<sup>1</sup>

1. O'Brien, T., Mihara, K., Sakaue, J., & Kimura, H. (2022). *Simply 500: Acing the TOEIC® Listening & Reading Test (Rev. ed.)*. Nan'un-do.<sup>1</sup>
2. Narisige, H. (2020). *TOEIC L&R Test Target 900* [TOEIC L&R Test, English grammar, Target 900]. J Research publishing.

The rationale for choosing these particular books stems firstly from their relatively recent

<sup>1</sup>The Educational Testing Service has disseminated official TOEIC materials aimed at catering to aspirants with target scores of 500, 650, and 800+. However, these resources offer a cursory representation for a diverse audience, making it arduous to discern the precise alignment of questions with specific score brackets. Consequently, I have refrained from incorporating these materials in my research.

publication dates, a factor that lends them an air of contemporary relevance. Additionally, the first tome has been explicitly designed as a textbook for educational purposes, complete with revision material, making it particularly well-suited for the intended audience. The second volume, meanwhile, contains multiple levels within its pages, enabling a more precise assessment of the target score in comparison to the use of multiple texts.

In this research, excerpts from questions listed at each proficiency level are evaluated by ChatGPT. I analyze how ChatGPT assesses these questions to determine whether it evaluates the difficulty of questions consistently across levels and distinguishes between difficulty levels based on the intended difficulty of the original questions.

Regarding the second research question, a dual-pronged approach to input is adopted: The initial component encompasses writing exemplars from a pre-intermediate learner, epitomizing the linguistic output germane to the target constituency (shown in Section 4.2); whereas the secondary input modality encompasses oral communication data excerpt from a TV program (details in Section 4.2), encapsulating a diverse range of communicative strategies.

The analysis of the first text aims to evaluate ChatGPT's ability to understand a written script that contains imperfections in terms of grammar and vocabulary usage. In contrast, the second text excerpts spoken language, which includes several applications of communication strategy. Additionally, this is an open source, allowing for further research by others.

## RESULTS AND DISCUSSION

### *Appraisal of Linguistic Intricacy*

This section scrutinizes how ChatGPT gauges the complexity of posed questions. The initial hurdle in this analysis is the insensitivity of ChatGPT to the difficulties that nonnative speakers are prone to encounter. Initially, I administered TOEIC questions with a score range of 500, 800, and 900 (each score approximately corresponds to CEFR levels A2 to B1, B2, and C1, respectively), encompassing ten questions per level. However, ChatGPT largely perceived them as facile without any specific guidance.

To circumvent this challenge, my subsequent attempt was to assess each question on a scale of 1–10. Nevertheless, the scores computed by ChatGPT were not static. Table 1 illustrates ten

attempts to evaluate the complexity of ten questions that are tailored for a 900 TOEIC score. This was achieved by giving a query to ChatGPT to rate the difficulty of each given question on a scale from 1 to 10. This reveals two salient features: firstly, ChatGPT still perceives TOEIC questions as straightforward, despite being significantly arduous for nonnative speakers aiming for a 900 score. Secondly, while there exist inconsistencies in the evaluations, the results tend to converge around a particular range, with standard deviations (SDs) of at most 1.5. As a result, ChatGPT's evaluation appears to have an identifiable threshold.

Table 1.

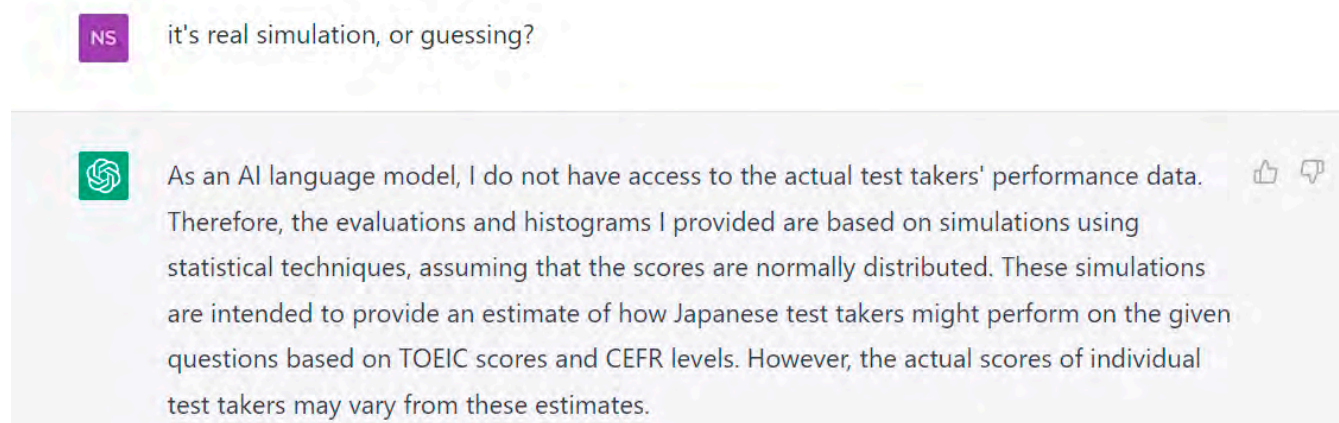
*Assessment of Complexity by ChatGPT*

Item No. /Trial	1	2	3	4	5	6	7	8	9	10	M	SD
1	3	2	2	5	3	2	3	3	3	5	3.1	1.044
2	4	3	3	6	6	3	2	2	5	6	4.0	1.550
3	2	4	4	4	4	4	5	4	6	3	4.0	1.0
4	4	5	5	4	2	5	4	4	5	4	4.2	0.871
5	3	3	6	7	4	6	5	3	4	5	4.6	1.356
6	2	2	3	3	2	3	4	2	4	3	2.8	0.748
7	3	2	2	4	2	2	3	3	4	4	2.9	0.831
8	2	2	2	2	2	2	2	2	3	2	2.1	0.3
9	5	6	5	7	2	6	5	5	6	7	5.4	1.356
10	2	2	3	3	3	3	3	3	5	4	3.1	0.831

In order to enhance the efficacy of the assessment, I conducted numerous iterations of simulations, amounting to 10,000 in total, (after performing calculations with a 1% margin of error, 95% confidence level, and assuming a population proportion of 50%, I arrived at a sample size of 9,601), and subsequently requested the provision of data pertaining to the mean and standard deviation. The authenticity of these simulations is corroborated by the ability of ChatGPT to furnish frequency distributions for each query, a capability that serves to confirm the veracity of the simulation in light of the fact that it does not possess access to the underlying datasets (as demonstrated in Figure 1). Thus far, this investigation has ascertained that the disclosed data is capable of reflecting the potential challenges associated with the evaluation, where the numerical values are mathematically convergent.

Figure 1.

Excerpt from ChatGPT about the Simulation and Data



In order to align the proficiency assessment with the specific characteristics of the TOEIC test, as perceived by the Japanese population (as an illustrative sample), I provided additional queries to facilitate a more refined evaluation. This approach aims to capture a more nuanced picture of the candidates' language capabilities under targeted conditions.

*Provide an assessment of the following inquiries for Japanese examinees utilizing the TOEIC scale ranging from 5 to 990, alongside the CEFR proficiency level. Please conduct 10,000 evaluations for each item, subsequently computing the mean and standard deviation based on the 10,000 scores. Subsequently, present the results in a tabular format as follows*

However, I conjectured that the modified assessment is fundamentally congruous with the conventional evaluation, with only minor variations in lexicon. In this investigation, a cohort of ten questions per tier, arbitrarily extracted from the respective texts, undergoes assessment by ChatGPT, with the outcomes delineated in Table 2.

The findings demonstrate a substantial disparity between the difficulty levels ascertained by ChatGPT and the intended scores, encompassing an extensive range. Notably, queries posited at the lowest stratum are appraised as the most arduous amongst the tripartite categorization. Consequently, it may be inferred that ChatGPT faces considerable challenges in evaluating TOEIC questions in a manner that accurately reflects the perceived difficulty for nonnative speakers.

Table 2.

ChatGPT's assessment of question difficulty for achieving specific scores

TARGET SCORE	500			800			900		
ITEM NUMBER	M	SD	CEFR	M	SD	CEFR	M	SD	CEFR
1	820	57	C1	722	122	C1	703	102	C1
2	825	49	C1	583	148	B2	783	91	C1
3	780	71	B2	846	75	C2	742	107	C1
4	830	51	C1	776	104	C1	790	94	C1
5	805	63	B2	745	115	C1	785	97	C1
6	855	38	C1	765	107	C1	741	108	C1
7	835	50	C1	718	125	C1	691	119	B2
8	800	70	B2	784	98	C1	562	130	B1
9	845	44	C1	767	106	C1	644	119	B2
10	820	58	C1	757	110	C1	689	119	B2
M of 10 items	822	22		746	68		713	72	

Of course, there remains a risk of “hallucination” wherein ChatGPT appears to perform tasks superficially without genuine processing. Nevertheless, it is evident that ChatGPT’s assessments lack consistency when measured against human evaluations, as its outputs represent what it deems the “most probable outcome” based on its extensive data training. This discrepancy underscores the inherent limitations in the model’s ability to mimic nuanced human judgment.

This analysis suggests that ChatGPT may not possess sufficient sensitivity to the difficulties that nonnative speakers encounter while responding to TOEIC questions. Even in traditional language acquisition models, Krashen’s Monitor Model posits that  $i+1$  is the optimal level of input for learners to effectively acquire a new language feature (1985). Even without invoking Krashen’s theory, it is widely agreed upon that language learning should proceed in a gradual and systematic manner. Therefore, an inability to discern subtle distinctions in the complexity of the English language may prove to be a disadvantage for learners.

Further research could be undertaken to enhance ChatGPT’s understanding of the challenges that nonnative speakers face, conceivably by assimilating more data from language learners or refining the model’s evaluation criteria.

#### **TOLERANCE OF LINGUISTIC COMPREHENSION IN NONNATIVE DISCOURSE**

This section delineates the capacity of ChatGPT to comprehend discourse produced by nonnative speakers, particularly in relation to grammatical proficiency and overall communicative competence. The primary focus is on the ability of ChatGPT to interpret the messages conveyed by a Japanese learner who is in the midst of linguistic development. To be more specific, the learner has achieved the 2nd grade in the Test of Practical English Proficiency (Eiken), which corresponds to a proficiency CEFR B1 (Eiken Foundation of Japan, n.d.).

*About nursery school.*

*I have a daughter, she is 4 years old recently do homework, remote work job. It's so busy, So I decide leave in nursery school, but it's very anxiety and nervousness. The truth is look after my kids is feeling*

*of relief but my job, housework, hanging out with friends. I have no choice leave my child at a day-care center. It is a first time leave my child there. Most of important thing is hows Nursery, what kind of nursery teacher is there. When I try leave my child at a day-care center, cultivate a good habit, cooperative personality, with another child.*

Based on Appendix A, the findings reveal that, despite the presence of various grammatical and lexical inaccuracies, ChatGPT is adept at discerning the intended message. As an extension of this investigation, I endeavor to identify the native language of the learner. In pursuit of this objective, a thorough analysis of the provided text is conducted for both versions, with version 4.0 offering a range of plausible predictions. To arrive at these conjectures, ChatGPT employs phonetic information derived from Japanese, and from a linguistic standpoint, demonstrates a commendable level of tolerance for inaccuracies in input. Moreover, the system exhibits an aptitude for applying linguistic knowledge pertinent to the respective languages. This attribute is particularly advantageous for employing ChatGPT in language learning contexts, as it inherently encompasses numerous linguistic discrepancies.

The ensuing investigation endeavors to evaluate the extent to which ChatGPT comprehends communication predicated not solely on linguistic constituents but also on extralinguistic dimensions, specifically sociolinguistic competence and strategic competence. In essence, though conventionally delineated, sociolinguistic competence entails the aptitude to efficaciously engage within a given context (Celce-Murcia et al., 1995), incorporating the interpretation of lexical constituents, norms, and discourse into sociocultural significance (Walcott, 2001). Strategic competence embodies the capacity to adroitly traverse scenarios wherein conversationalists confront obstacles in conveying their intended meaning while employing the target language (Savignon, 1997), encompassing tactics such as circumlocution, paraphrasing, linguistic appropriation, entreaty for assistance, and gesticulation (Tarone, 1984).

Subsequent extracts (Appendix B) have been procured from an erudite gastronomic discourse

showcased on the distinguished Japanese television production, “King & Prince-ru” (broadcasted by Nippon TV). One of the broadcasts highlights the illustrious Japanese idol group, King & Prince endeavoring to perform a recipe in English. Although the program’s intent is to evoke subtle mirth among the audience through the occasional infelicitous or incongruous utilization of English by Japanese speakers, I perceive it as a quintessential exemplification of sociolinguistic and strategic competence necessary for achieving a demanding task in English, for both speakers and interlocutors. Significantly, the terminology of edibles poses a daunting aspect, as a multitude of Japanese individuals may be unacquainted with the precise designations.

The passage exemplifies the initial stage of culinary preparation, encompassing ingredient organization. The instruction delineated, “Slice the onion thinly, cut the carrot into thin strips, and chop the shungiku (crown daisy) into 5cm pieces. Quarter the scallops.” The speaker subsequently implored:

*Ok, listen, Can you bring onion? Onion is thin cut, Carrot, slim cut, part 3, spring leaves are used when we eat Sukiyaki, cut 5 cm, and finally, Little mermaid brassiere.*

The speaker wrestled with the lexemes shungiku and scallops, endeavoring to communicate them indirectly. Shungiku is denoted as “ ” (literally translated as spring-chrysanthemum), and the speaker employed the term “spring” and the concept of “leaves” to depict the vegetable. The speaker also proffered ancillary information that it is employed in Sukiyaki, a Japanese hot pot. To elucidate scallops, the speaker resorted to invoking an image of Ariel, the Little Mermaid, donning a brassiere composed of scallops as a supplementary referent to compensate for his limited lexicon. Although the interlocutors encountered several obstacles, they ultimately deciphered the message and procured the relevant food ingredients.

I presented a comparable instruction to ChatGPT versions 3.5 and 4, with the outcomes revealing that both iterations could readily interpret the initial two items but faced challenges in discerning the third and fourth. Version 3.5 could only offer several alternative probabilities for “spring leaves” but failed to deduce the accurate

response, even with additional information provided. In contrast, ChatGPT 4 ultimately identified “shungiku” by leveraging the contextual cue that it is utilized in Sukiyaki. This outcome suggests that ChatGPT can incorporate certain cultural notions in its predictions but does not strive to exploit linguistic correlations derived from the target learners’ native language.

The fourth item presented the most substantial challenge, with version 3.5 conceding relatively early and requesting further information, while version 4 managed to extract an oblique expression of the intended term. However, even with additional guidance that the ingredient is seafood, version 4 could only propose several potential alternatives.

In consideration of these excerpts, it can be inferred that ChatGPT exhibits a degree of flexibility in interpreting nonnative speakers’ discourse when strong contextual pertinence supports the intended message. Nevertheless, it may be impractical to expect the accurate response only using contrasting language cues. As a result, human interlocutors, particularly those proficient in accommodating nonnative speakers, possess an elevated probability of comprehending such discourse and nurturing motivation through effective communication.

In this regard, for further research, it is advisable to conduct a comparative analysis between human instructors and ChatGPT, particularly focusing on the significance of nonnative speakers in language teaching.

## CONCLUSION

This investigation endeavors to scrutinize ChatGPT’s aptitude in deciphering input from nonnative interlocutors, specifically focusing on those acquiring English as a second language, with Japanese speakers representing the archetypal population. The empirical analysis examines ChatGPT’s assessment of difficulty gradients within the TOEIC. Furthermore, this study quantifies ChatGPT’s responsiveness to discourses generated by nonnative speakers, encompassing both syntactically flawed utterances and those integrating sociolinguistic or strategic competence expressions.

The results reveal that ChatGPT exhibits dexterity in lexical and grammatical constituents, while maintaining the capacity to decipher imperfect

syntactic discourses elicited by nonnative speakers. Nonetheless, ChatGPT fails to accurately discern the fluctuating degrees of arduousness experienced by nonnative interlocutors. Additionally, although ChatGPT demonstrates comprehension of socio-linguistic expressions when contextual cues are unambiguous, its grasp of alternative communication tactics employed by nonnative speakers remains relatively circumscribed.

The findings of this inquiry attest to ChatGPT's linguistic proficiency being highly pertinent in the context of apprehending nonnative linguistic variations. However, the model exhibits shortcomings in addressing nonlinguistic factors. Consequently, human instructors retain an indispensable role in language acquisition, given their capacity to comprehend the wider array of communicative channels utilized by learners. ChatGPT may be perceived as a language model analogous to a native speaker, thereby presenting potential obstacles for nonnative speakers, particularly those with limited proficiency, when engaging in output activities. Such impediments may stem from the learners' linguistic constraints or their inability to apprehend the strategies they employ.

The study's limitations arise from its dependency on a circumscribed data set; hence, supplementary research could enrich our comprehension of ChatGPT's capabilities. Notwithstanding, the proffered data cogently elucidates diverse communicative facets, serving as a preliminary benchmark towards integrating this avant-garde technology into language acquisition.

The advent of sophisticated technology, capable of managing an expansive array of tasks, persistently incites contentious deliberations, often accompanied by apprehensions concerning adverse ramifications. Nevertheless, akin to other domains, ChatGPT can be harnessed positively when wielded by adroit educators who possess the requisite experience and acumen to unleash its full potential.

Disclosure statement: The author reports there are no competing interests to declare.

# References

- Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, 2(1), 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Árva, V., & Medgyes, P. (2000). Native and non-native teachers in the classroom. *System*, 28(3), 355–372. [https://doi.org/10.1016/S0346-251X\(00\)00017-8](https://doi.org/10.1016/S0346-251X(00)00017-8)
- Calafato, R. (2019). The non-native speaker teacher as proficient multilingual: A critical review of research from 2009–2018. *Lingua*, 227, 102700. <https://doi.org/10.1016/j.lingua.2019.06.001>
- Celce-Murcia, M., Dornyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6(2), 5–35. <https://doi.org/10.5070/L462005216>
- Cheng, C.-C. (2020). Game-based learning and TOEIC vocabulary building: NNEs students' learning attitudes and motivations. *Journal of Applied Linguistics and Language Research*, 7(1), 135–143. <https://www.jallr.com/index.php/JALLR/article/view/1085>
- Clark, E., & Paran, A. (2007). The employability of non-native-speaker teachers of EFL: A UK survey. *System*, 35(4), 407–430.
- Eiken Foundation of Japan. (n.d.). Details of the EIKEN CSE Score. Retrieved from [https://www.eiken.or.jp/eiken/en/eiken-tests/overview/eiken-cse\\_admission.html](https://www.eiken.or.jp/eiken/en/eiken-tests/overview/eiken-cse_admission.html)
- Floris, D. F., & Renandya, W. A. (2020). Promoting the value of non-native English-speaking teachers. *PASAA*, 59(1), Article 1. <https://doi.org/10.58837/CHULA.PASAA.59.1.1>
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected Readings* (pp. 269–293). Penguin.
- Kemaloglu-Er, E., & Lowe, R. J. (2022). Language teacher identity, world Englishes, and ELF: A duoethnography between a “native speaker” teacher and a “non-native speaker” teacher. *Journal of Language, Identity & Education*, 22(5), 478–491. <https://doi.org/10.1080/15348458.2022.2141244>
- Kim, N.-Y. (2022). English with AI: A new era of TOEIC learning for students majoring in airline services. *Linguistic Research*, 39(Special edition), 97–122. <https://doi.org/10.17250/kh-isli.39..202209.004>
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Addison-Wesley.
- Lasagabaster, D., & Sierra, J. M. (2002) University students' perceptions of native and non-native speaker teachers of English. *Language Awareness*, 11(2), 132–142. <https://doi.org/10.1080/09658410208667051>
- Mahboob, A. (2004). Native or non-native: What do the students think? In L. D. Kamhi-Stein (Ed.), *Learning and teaching from experience: Perspectives on nonnative English-speaking professionals* (pp. 1–31). University of Michigan Press.
- Matassoni, M., Gretter, R. Falavigna, D., & Giuliani, D. Non-native children speech recognition through transfer learning. (2018). In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, Canada, (pp. 6229–6233). <https://doi.org/10.1109/ICASSP.2018.8462059>
- Matsuda, A., & Matsuda, P. K. (2001). Autonomy and collaboration in teacher education: Journal sharing among native and nonnative English-speaking teachers. *CATESOL Journal*, 13(1) 109–121. [http://www.catesoljournal.org/wp-content/uploads/2014/07/CJ13\\_matsuda.pdf](http://www.catesoljournal.org/wp-content/uploads/2014/07/CJ13_matsuda.pdf)
- Miyake, K. (2007). How young Japanese express their emotions visually in mobile phone messages: A sociolinguistic analysis. *Japanese Studies*, 27(1), 53–72. <https://doi.org/10.1080/10371390701268646>
- Narisige, H. (2020). *TOEIC L&R Test Target 900 [TOEIC L&R Test, English grammar, Target 900]*. J Research Publishing.
- Nippon TV. (2022, November 19). King & Prince & ( ) [King & Prince: No Japanese Allowed Cooking & Nagase's (Secret) Apprenticeship with a Professional Cleaner].
- O'Brien, T., Mihara, K., Sakaue, J., & Kimura, H. (2022). *Simply 500: Acing the TOEIC listening & reading test (Rev. Edition)*. Nan'un-do.
- Open AI (2023). ChatGpt. <https://openai.com/blog/chatgpt>
- Ozawa, S. (2019). Effects of Japanese university students' characteristics on the use of an online English course and TOEIC scores. *CALICO Journal*, 36(3), 225–239.
- Sakai, N. (2007). The theoretical study of motivational transfer and entertainment use in self study CALL. *Asian EFL Journal*, Thesis section. [https://www.asian-efl-journal.com/Thesis\\_Sakai.pdf](https://www.asian-efl-journal.com/Thesis_Sakai.pdf)
- Sakai, N. (2013). The role of sentence closing as emotional markers: A case of Japanese mobile phone e-mail. *Discourse, Context and Media*, 2(3) 149-155, <https://doi.org/10.1016/j.dcm.2013.07.001>
- Sakai, N. (2015). A study of translating extra-textual expressions from non-English language into English: The case of contemporary Japanese computer-mediated communication. In Leong, K, and Chen, P, (Eds.). *Translation and Cross-cultural Communication Studies in the Asia Pacific*. Leiden, Netherlands: Brill. 227-238. <http://dx.doi.org/10.1163/9789004299245>
- Sakai, N. (2023). Investigating the Feasibility of ChatGPT for

Personalized English Language Learning: A Case Study on its Applicability to Japanese Students. The Journal of Okayama e-learning Research Group, 1, 29-49

Savignon, S. J. (1997). *Communicative competence: Theory and classroom practice: Texts and Contexts in Second Language Learning* (2nd ed.). McGraw Hill.

SciencePo. (2023, September 2). ChatGPT: Sciences Po implements rules and opens up discussion about AI in higher education. SciencesPo. <https://www.sciencespo.fr/en/news/sciences-po-implements-strict-rules-about-the-use-of-chatgpt-by-students>

Selvi, A. F. (2011). The non-native speaker teacher. *ELT Journal*, 65(2), 187–189. <https://doi.org/10.1093/elt/ccq092>

Tarone, E. (1984). Teaching strategic competence in the foreign-language classroom. In S. J. Savignon & M. S. Berns (Eds.), *Initiatives in communicative language teaching: A book of readings* (pp. 127–136). Addison Wesley.

Todd, R. W., & Pojanapunya, P. (2009). Implicit attitudes towards native and non-native speaker teachers. *System*, 37(1), 23–33. <https://doi.org/10.1016/j.system.2008.08.002>

Vera, K., & Laurence, D.-H. (2022) Machine translation: Friend or foe in the language classroom? *L2 Journal: An Open Access Refereed Journal for World Language Educators*, 14(1), 105–135. <https://doi.org/10.5070/L214151723>

Walcott, W. H. (2001). Knowledge, competence, and communication [Electronic Version]. *Radical Pedagogy*, 3(3). [http://radicalpedagogy.icaap.org/content/issue3\\_3/2-walcott.html](http://radicalpedagogy.icaap.org/content/issue3_3/2-walcott.html)

Wijayanti, P. R., Suwartono, T., & Kusuma, L. (2022). Perception of vocational school students towards online learning and their TOEIC scores. *International Journal for Educational and Vocational Studies*, 4(1), 35–41. <https://doi.org/10.29103/ijevs.v4i1.5734>

## APPENDIX A1

### FULL EXCERPT FROM CHATGPT FOR GRAMMATICAL UNDERSTANDING (CHATGPT 3.5)

NS

Can you understand the following?



About nursery school.

I have a dauter, she is 4 years old recently do homework, remote work job. It's so busy, So I decide leave in nursery school, but it's very anxiety and nervousness. The truth is look after my kids is feeling of relief but my job, hosework, hanging out wiz friends. I have no choice leave my child at a day-care center. It is a first time leave my child there. Most of important thing is hows Nursey, what kind of nusery teacher is there. When I try leave my child at a day-care center, cultivate a good habit, cooperative personality, with another child.



Yes, I understand the message. It seems like the person who wrote this is a parent with a 4-year-old daughter who is struggling to balance their work and home responsibilities. They are considering leaving their daughter at a nursery school for the first time, but are feeling anxious and nervous about it. The writer believes that leaving their child at a nursery school may be necessary to allow them to focus on their job, household chores, and social life, but they are concerned about the quality of care their child will receive and the influence of the nursery teachers and other children on their daughter's habits and personality.



NS

Can you predict what is the native language of this writing?



Based on the grammar and phrasing used in the writing, it is difficult to determine the writer's native language with certainty. However, there are some clues in the use of certain phrases that suggest the writer may not be a native English speaker. For example, "It's so busy, So I decide leave in nursery school" may indicate that the writer's native language does not have the same word order as English. Additionally, the use of "hosework" instead of "housework" and "nusery" instead of "nursery" may suggest that the writer's native language is one where those letters or sounds do not exist or are pronounced differently.



NS

Can you say several possibilities?



Sure, based on the clues in the writing, here are several possibilities for the writer's native language:



- A Romance language such as Spanish, French, or Italian, where the word order is different from English and where the writer may be translating phrases directly from their native language.
- An Asian language such as Chinese, Japanese, or Korean, where the writer may have difficulty with English spelling and pronunciation due to differences in the phonetic system.
- A Slavic language such as Russian, Polish, or Czech, where the writer may have difficulty with certain English sounds and word order due to differences in the grammar and pronunciation of their native language.



More specific?



While it is difficult to determine the writer's native language with certainty based on the limited sample of text provided, here are some additional clues that could help narrow down the possibilities:



- The writer uses the phrase "do homework" instead of "complete homework," which suggests that their native language may not use the same verb for this activity.
- The writer uses the phrase "feeling of relief" instead of "sense of relief," which could be a direct translation from their native language.
- The writer uses the phrase "hang out wiz friends" instead of "hang out with friends," which may suggest that their native language does not have the "th" sound or that they are not accustomed to writing in English.

Based on these additional clues, it is possible that the writer's native language could be:

- Spanish, as the phrase "do homework" could be a direct translation of "hacer tarea," and the use of "wiz" instead of "with" may be due to the similar pronunciation of "z" and "th" in some Spanish dialects.
- Chinese, as the use of "wiz" instead of "with" may be due to the difficulty Chinese speakers often have with the "th" sound, and the phrase "feeling of relief" may be a direct translation of a similar phrase in Chinese.
- Polish, as the writer uses "hosework" instead of "housework," which could be a direct translation of the Polish word "praca domowa," and the use of "wiz" instead of "with" could be due to the similar pronunciation of the Polish letter "ż" and the English "z."

Again, it's important to note that these are only possibilities and should not be taken as definitive conclusions. Without more context or information, it is not possible to accurately predict the writer's native language.

## APPENDIX A2

### FULL EXCERPT FROM CHATGPT FOR GRAMMATICAL UNDERSTANDING (CHATGPT 4.0)

## APPENDIX B1

### FULL EXCERPT FROM CHATGPT FOR SOCIOLINGUISTIC UNDERSTANDING (CHATGPT 3.5)

NS Do you understand the following by Japanese learner:

Ok, listen, Can you bring onion? Onion is thin cut, Carrot, slim cut, part 3, spring leaves is when we eat Sukiyaki, cut 5 cm, then, Little mermaid brassiere.



I understand that this text is written by a Japanese learner and may contain some grammatical errors or unconventional phrasing. Here is my interpretation of the text:



The person is asking someone to bring the following ingredients for a dish, possibly Sukiyaki:

1. Thinly cut onions
2. Slimly cut carrots
3. Spring leaves, cut into 5 cm pieces

However, the phrase "part 3" is unclear, and it's not obvious how it relates to the other items. Additionally, the mention of "Little mermaid brassiere" seems unrelated to the context of cooking, and I am unsure how to interpret it in this situation.

NS

The person wants to tell how to cook something



Based on the given text, the person appears to be providing instructions on how to prepare ingredients for a dish, likely Sukiyaki. Here is a revised interpretation of the instructions:



1. Thinly slice onions.
2. Cut carrots into slim strips.
3. Cut spring leaves into 5 cm pieces.

The person seems to be describing the desired size and shape for each ingredient before cooking. However, the phrases "part 3" and "Little mermaid brassiere" remain unclear, and it is difficult to understand how they fit into the context of cooking instructions.

NS

"Little mermaid brassiere" indicates some sea food...

## APPENDIX B2

### FULL EXCERPT FROM CHATGPT FOR SOCIOLINGUISTIC UNDERSTANDING (CHATGPT 4.0)



I see, in that case, the mention of "Little mermaid brassiere" could be a creative way of referring to seafood, perhaps inspired by the character Ariel from Disney's "The Little Mermaid." However, without more context or clarification, it is difficult to determine the specific type of seafood the person is referring to or how it should be prepared in the dish.



What would be spring leaves



"Spring leaves" typically refers to the young, tender leaves of various plants and herbs that emerge during the spring season. These leaves are often used in cooking for their fresh and delicate flavors. Examples of spring leaves include baby spinach, young kale leaves, arugula, and various types of lettuce. In the context of the original text, it seems the person is instructing someone to cut spring leaves into 5 cm pieces, possibly for use in a Sukiyaki dish.



What is the exact name of vegetable the person would like to indicate?



It's difficult to determine the exact vegetable the person is referring to when they mention "spring leaves," as the term is quite broad and can encompass various types of young, tender leaves. However, considering the context of a Sukiyaki dish, the person might be referring to "shungiku" or "chrysanthemum greens," which are a common ingredient in Japanese cuisine and often used in Sukiyaki. Shungiku has a unique, slightly bitter flavor and adds a nice contrast to the other ingredients in the dish.



Um, "Little mermaid brassiere"... please tell me some possible materials