Съпоставителен анализ между автентични материали и еквиваленти, генерирани от изкуствен интелект

Евгени Николаев Станчев

Comparative analysis between authentic studying materials and AI generated equivalents for level A2

Evgeni Nikolaev Stanchev

Abstract:
This paper focuses on implementing and using computer generated texts, to assist English as a Second Language (ESL) teachers when choosing additional teaching materials. A study was done, involving analyzing all the authentic written discourse in the student book "Speak out" 2nd edition, published by Pearson / Longman, with authors Frances Eals and Steve Oaks. All the texts in the book (49) were provided to an artificial intelligence (AI) program – ChatGPT. The AI was instructed to write a similar story, while bearing in mind the intended level of the students – A2. Two corpuses were then created – one for the authentic texts, and one for the AI generated ones. The results show that the authentic texts had an average text length of 121.16 words per text, while the AI generated had an average text length of 145.1. In conclusion, with every numerical criterion that was tested, the results showed that AI was adequately equipped with the necessary tools to write clear and concise texts, which could be used to assist ESL teachers as a source of extra or main studying materials.
Keywords: ESL, corpus analysis, AI, ChatGPT

For contacts: Evgeni Stanchev, evgeni.stanchev@mu-varna.bg

INTRODUCTION

In recent years, there has been a growing interest in the application of artificial intelligence (AI) technologies in education (Gibson, 2019; Johnson et al., 2021). Among these technologies, conversational agents, or chatbots, have gained attention for their potential to enhance learning experiences (Wang & Heffernan, 2018). Chatbots, such as ChatGPT, are AI systems capable of engaging in human-like conversations and providing personalized assistance across various domains (Radanovic et al., 2020). In the context of education, chatbots have been proposed as tools to support student learning, facilitate communication, and provide instant feedback (Baylor & Kim, 2018).

Methods

Data collection

The collected data from two sources: (1) all the authentically written texts in the textbook "Speak Out" 2nd Edition – Elementary, by Pearson / Longman, with authors Frances Eals and Steve Oaks, and (2) text passages generated by ChatGPT. For the ChatGPT-generated text, OpenAI GPT-3.5 model was utilized to generated passages related to the exact educational topics. The AI was instructed to create a text, on the same topic and with a similar length, with a high focus on the educational level – A2 – Elementary. A total of 49 texts were generated to mimic the style and content typically found in a student book, which covered all the authentic texts.

The authentic texts were 49, with a total word count of 5937 words. This is an average of 121.16 words per text. The ChatGPT-generated text had a total word count of 7110 with an average of 145.1 words per text.

Text analysis

For linguistic analysis of the ChatGPT-generated text and the student book corpus, natural language processing (NLP) techniques were employed. Initially, the text passages were tokenized and performed part-of-speech tagging. Additionally, to explore syntactic patterns and words usage, Sketch Engine was utilized (Kilgarriff et al., 2004), a robust tool for corpus linguistics. Sketch engine was capable to identify collocations, phrase structures, and semantic relationships within the text, providing valuable insights into the linguistic features of the passages.

Results

The authentic texts showed a total of 54 n-grams 2-grams, with an absolute frequency of 449 and a total of 4 n-gram 3-grams, with an absolute frequency of 26.

54 items | 449 frequency	n-gram 2-grams. word	in the (43)	of the (20)	the world (17)	in a (16)	on the (13)	to the (13)	in your (11)	the most (10)	did not (10)	is the (10)
4 items | 26 frequency	n-gram 3-grams. word	in the world (11)	in your bag (5)	do you do (5)	around the world (5)						

In comparison, the ChatGPT-generated text had a total of 35 n-gram 2-grams, with an absolute frequency of 323 and only 1 n-gram 3-grams with an absolute frequency of 5. These were the top 10 used.

35 items | 323 frequency	n-gram 2-grams. word	in the (41)	on the (17)	of the (17)	to the (17)	in a (15)	with a (14)	in your (11)	from the (10)	into the (10)	the world (10)
1 item | 5 frequency	n-gram 3-grams. word	in your bag (5)									

The authentic texts had a total of 183 adjectives, with an absolute frequency of 453. These were the top 10 used.

word list - adjective	good (20)	big (16)	new (14)	more (11)	different (11)	small (10)	late (10)	other (9)	beautiful (9)	important (9)

The ChatGPT-generated texts had nearly double the number, with 348 adjectives with an absolute frequency of 596. These were the top 10 used.

word list - adjective	new (14)	significant (9)	such (9)	small (8)	late (8)	unique (8)	other (7)	local (7)	perfect (6)	next (6)

The authentic texts had a total of 71 adverbs with an absolute frequency of 304.

not (47)	so (20)	very (18)	then (16)	only (13)	too (11)	usually (10)	most (9)	also (9)	often (8)

The ChatGPT-generated texts had more than double the amount with 148 adverbs with an absolute frequency of 327. These were the top 10 used.

not (28)	most (13)	however (11)	only (11)	more (8)	just (7)	then (7)	well (7)	often (6)	here (6)

The authentic texts had a total of 3 conjunctions with an absolute frequency of 334.

and (256)	but (46)	or (32)

In comparison, the ChatGPT-generated text had 4 conjunctions with an absolute frequency of 269.

and (221)	or (27)	but (20)	nor (1)

The authentic texts had a total lemma count of 1221 items, with an absolute frequency of 7028. These were the top 10 used.

. (374)	the (293)	, (283)	and (258)	be (251)	a (213)	in (180)	to (122)	of (108)	you (87)

In comparison, the ChatGPT-generated text had 2068 items with n total frequency of 8438. These were the top 10 used.

, (516)	. (383)	the (327)	a (262)	and (224)	to (192)	in (185)	be (169)	of (159)	for (83)

The authentic texts contained 760 different nouns, with a total frequency of 1765. These were the top 10 used.

people (32)	time (29)	city (23)	year (20)	friend (20)	world (18)	place (17)	family (15)	car (15)	child (13)

In contrast, the ChatGPT-generated texts had 1196 items with a total frequency of 2251. These were the top 10 used.

city (22)	family (19)	world (18)	time (17)	year (16)	car (16)	friend (16)	festival (14)	life (13)	journey (13)

The authentic texts contained a total of 20 items and a frequency of 121 for numerals. These were the top 10 used.

[number] (35)	one (26)	two (17)	three (6)	fifteen (5)	five (4)	fifty (3)	seven (3)	sixteen (3)	ten (3)

In comparison, the ChatGPT-generated texts contained also 20 items and a total frequency of 103. These were the top 10 used.

[number] (30)	one (19)	two (11)	three (6)	fifteen (5)	five (3)	seven (3)	sixteen (3)	fifty (3)	ten (3)

The authentic texts contained a total of 38 items with a total frequency of 703 for prepositions. These were the top 10 used.

in (180)	of (108)	for (66)	to (47)	on (47)	with (39)	at (37)	about (23)	from (18)	by (15)

In comparison, the ChatGPT-generated texts contained 55 items with a total frequency of 977 for prepositions. These were the top 10 used.

in (185)	of (159)	for (83)	to (73)	with (68)	on (60)	as (45)	from (41)	at (32)	by (24)

The authentic texts contain a total of 19 items with a total frequency of 575 for pronouns. These were the top 10 used.

you (87)	it (76)	i (68)	we (48)	she (45)	they (42)	her (41)	your (34)	my (31)	our (18)

The ChatGPT-generated texts contained 24 items with a total frequency of 551 for pronouns. These were the top 10 used.

you (63)	it (54)	your (52)	i (48)	her (46)	we (46)	my (37)	our (37)	their (28)	she (27)

The Authentic texts contained a total of 158 items with a total frequency of 967 for verbs. These were the top 10 used.

be (251)	do (68)	have (55)	go (36)	get (27)	say (19)	see (17)	find (16)	eat (16)	work (15)

The ChatGPT-generated texts contained 451 items with a total frequency of 1156 for verbs. These were the top 10 used.

be (169)	have (34)	do (27)	explore (16)	take (14)	find (13)	offer (12)	join (12)	go (11)	follow (10)

The last token analyzed by the corpus was the word token. The authentic texts contained 1506 items and a total frequency of 7028. These were the top 10 used.

. (374)	the (293)	, (283)	and (258)	a (188)	in (180)	to (122)	of (108)	you (87)	it (76)

The ChatGPT-generated texts contained a total of 2495 items with a total frequency of 8438 for the word token. These were the top 10 used.

, (516)	. (383)	the (327)	a (232)	and (224)	to (192)	in (185)	of (159)	for (83)	with (68)

CONCLUSION

In conclusion, the results of this study provide valuable insights regarding the potential of chatbot-generated content in educational contexts. Our analysis revealed that while there are differences in syntactic structures and quality between ChatGPT-

generated text and what is traditionally accepted as educational system content, ChatGPT demonstrates superiority in several key aspects.

Firstly, ChatGPT-generated text exhibited higher readability scores compared to traditional ESL authentically created texts, suggesting that it may offer a more accessible and comprehensible learning experience for students across different proficiency levels. The simplified sentence structures and clear language used in ChatGPT-generated content have the potential to enhance student comprehension and engagement, particularly in subjects with complex concepts or technical terminology.

Overall, the flexibility and scalability of ChatGPT allow for rapid content generations and customization, addressing the growing demand for adaptable educational resources in today's digital age. Educators can leverage ChatGPT to create personalized learning materials tailored to individual student needs and preferences, and in doing so fostering a more inclusive and student-centered learning environment.

REFERENCE

Baylor, A. L., & Kim, Y. (2018). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. Journal of Computer Assisted Learning, 34(6), 501-509.

Gibson, A. (2019). Artificial intelligence in education: Promises and challenges. Educational Media International, 56(2), 97-112.

Johnson, J. R., Haigh, K. Z., & Gifford, T. (2021). The use of artificial intelligence in medical education: A review. Medical Education Online, 26(1), 1842163.

Kilgarriff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The Sketch Engine. Information Technology & Lexicography, 2, 105-116.

Radanovic, G., Damper, R. I., Higgins, C. A., & Eder, S. A. (2020). The potential of chatbots in education: A systematic review. Computers & Education, 148, 103819.

Wang, Y., & Heffernan, N. (2018). Investigating the impact of personalized tutoring with state assignment features in a Logic Tutor. Journal of Educational Data Mining, 10(1), 46-75.