



An Analysis of Syntactic Complexity in Texts Translated by Google Translate and Human Translator from Arabic into English

Aziz Mohammed Abdo Saeed^{*}

aziz.saeed@hotmail.com

Abstract:

This paper compared the syntactic complexity of translations from Arabic into English produced by Google Translate with human translation of the same texts using the web-based L2 Syntactic Complexity Analyzer (L2SCA). Syntactic complexity is an important aspect of writing quality and language proficiency in general. Understanding how machine translation tools deal with syntactic complexity when translating texts from Arabic into English will offer insights for their efficacy. The study analyzed the syntactic complexity of translations of five text samples collected from an Arabic book "Hiwar Ma'a Sadeeqi Almulhid" [*Dialogue with An Atheist*] by Mostafa Mahmoud. The output of Google Translate is compared with human translation for the same texts by Mohammed Yahia. The results showed that Google Translate output has lower syntactic complexity ratios for most of the 10 indices measured. The paper highlighted key findings that have implications for developing machine translation quality. This paper contributes to the ongoing development of machine translation.

Keywords: Syntactic Complexity, Human Translation, Machine Translation, Google Translate.

^{*} Master of English Linguistics, English Department, Faculty of Arts, Thamar University, Republic of Yemen.

Cite this article as: Saeed, A. M. A. (2025). An Analysis of Syntactic Complexity in Texts Translated by Google Translate and Human Translator from Arabic into English, *Arts for Linguistic & Literary Studies*, 7(1): 694 -710.

© This material is published under the license of Attribution 4.0 International (CC BY 4.0), which allows the user to copy and redistribute the material in any medium or format. It also allows adapting, transforming or adding to the material for any purpose, even commercially, as long as such modifications are highlighted and the material is credited to its author.



تحليل التعقيد النحوي في نصوص مترجمة من العربية إلى الإنجليزية: مقارنة ترجمة جوجل بالترجمة البشرية

عزیز محمد عبده سعيد *

aziz.saeed@hotmail.com

ملخص:

قارنت هذه الدراسة بين التعقيد النحوي للترجمات من العربية إلى الإنجليزية التي ينتجها موقع جوجل للترجمة، والترجمة البشرية للنصوص نفسها باستخدام محلل التعقيد النحوي L2SCA الآلي. ويعد التعقيد النحوي جانباً مهماً من جودة الكتابة وكفاءة استخدام اللغة بشكل عام. إن فهم كيفية تعامل أدوات الترجمة الآلية مع التعقيد النحوي عند ترجمة النصوص من العربية إلى الإنجليزية سيوفر رؤى حول فعالية هذه الأدوات. وقد قامت هذه الدراسة بتحليل التعقيد النحوي لترجمات خمس عينات نصية تم جمعها من كتاب عربي بعنوان (حوار مع صديقي الملحد) للكاتب مصطفى محمود، وتم مقارنة مخرجات ترجمة جوجل بالترجمة البشرية للنصوص نفسها لمحمد يحيى. وقد أظهرت النتائج أن مخرجات ترجمة جوجل لها نسب تعقيد نحوي أقل بالنسبة إلى معظم المؤشرات العشرة التي تم قياسها. وقد تم تسليط الضوء على أهم النتائج وأثرها في تطوير جودة الترجمة الآلية، مع وضع توصيات لدراسات مستقبلية لتطوير جودة الترجمة الآلية. وبذلك فإن هذا البحث يسهم في التطوير المستمر للترجمة الآلية.

الكلمات المفتاحية: التعقيد النحوي، الترجمة البشرية، الترجمة الآلية، ترجمة جوجل.

* ماجستير لغويات - قسم اللغة الإنجليزية - كلية الآداب - جامعة ذمار - الجمهورية اليمنية.

للاقتباس: سعيد، ع. م. ع. (2025). تحليل التعقيد النحوي في نصوص مترجمة من العربية إلى الإنجليزية: مقارنة ترجمة جوجل بالترجمة البشرية، *الآداب للدراسات اللغوية والأدبية*، 7(1): 694-710.

© نُشر هذا البحث وفقاً لشروط الرخصة Attribution 4.0 International (CC BY 4.0)، التي تسمح بنسخ البحث وتوزيعه ونقله بأي شكل من الأشكال، كما تسمح بتكييف البحث أو تحويله أو إضافته إليه لأي غرض كان، بما في ذلك الأغراض التجارية، شريطة نسبة العمل إلى صاحبه مع بيان أي تعديلات أجريت عليه.



1- Introduction

In the era of globalization, the demand for effective communication across languages has surged, leading to an increased reliance on translation services. Among these, machine translation (MT) systems, particularly Google Translate, have gained significant popularity due to their accessibility and rapid processing capabilities. However, the quality of translations produced by MT systems often raises questions regarding their quality compared to human translations. Syntactic complexity is one of the aspects of machine translation quality. It is defined as the range and sophistication of sentence structures used in a text, and it plays a crucial role in determining the overall quality and readability of translated content (Ortega, 2003; Ha, 2023). It also plays a crucial role in determining writing proficiency level (Park, 2012; Yang et al., 2015; Yu, 2019; Zeng, 2023).

One of the key requirements for academic success has been identified as writing ability. Internationally known standardized tests like the TOEFL and IELTS, which gauge English language learners' proficiency in essay writing, typically measure writing. Scholars and practitioners of second language writing have given the problem of syntactic complexity in writing and its relationship to writing proficiency a great deal of attention. The variety of language patterns and their level of sophistication that are produced in spoken or written conversation are commonly thought of as syntactic complexity (Ortega, 2003). Additionally, the importance of syntactic complexity in second language writing has been highlighted by numerous previous researches that examined the close relationship between syntactic complexity in L2 writing and language proficiency or the caliber of L2 writing during the decade (Cancino & Panes, 2021; Lu, 2011; Ortega, 2003).

As such, understanding the differences in syntactic complexity between translations generated by Google Translate and those produced by human translators is essential for evaluating the effectiveness of MT systems in conveying syntactic structures and maintaining the integrity of the source text. Syntactic complexity measures are essential research instruments in a variety of language-related fields, including second language acquisition (SLA). The syntactic complexity must be viewed as a multifaceted construct and evaluated at several structural levels, such as sentence, clausal, and phrasal levels (Ha, 2023). The difference in vocabulary, syntax, sounds and style presents challenges for the translation between the two languages (Alnasery, 2024).

This research aims to explore these differences by conducting a comparative analysis of syntactic complexity in translations produced by Google Translate with translations by human translators for the same texts. By employing a multidimensional approach to syntactic complexity, this study will assess various measures, including length of production unit, amount of coordination, amount of subordination and level of



phrasal sophistication, to provide a comprehensive understanding of how each translation method handles syntactic elaboration.

The findings of this study will contribute to the ongoing discourse on the efficacy of machine translation in professional and academic settings, offering insights into the potential advantages and limitations of relying on automated systems for translation tasks. Furthermore, this research will inform educators and practitioners about the implications of using MT tools in language learning and translation practices, ultimately enhancing the quality of translated texts in an increasingly interconnected world.

In summary, this study seeks to answer the following research questions:

1. How do the measures of syntactic complexity compare between human and Google Translate translations?
2. How do the differences in syntactic complexity levels serve to recognize translations generated by Google Translate?

By addressing these questions, this research will shed light on the capabilities and shortcomings of Google Translate in the context of syntactic complexity, providing valuable insights for both researchers and practitioners in the field of translation studies.

2- Literature Review

Syntactic complexity is a multifaceted construct that refers to the range and sophistication of grammatical structures used in language production. It has been widely studied in the context of both first language (L1) and second language (L2) writing, with significant implications for assessing writing quality and proficiency (Hwang et al., 2020; Ortega, 2003; Norris & Ortega, 2009). Various measures have been proposed to quantify syntactic complexity, including the mean length of sentences, the number of clauses per sentence, and the use of complex structures such as subordinate clauses (Lu, 2011; Bulté & Housen, 2014). These measures are crucial for understanding how syntactic complexity correlates with writing proficiency, as more complex syntactic structures are often indicative of higher writing skills (Kisselev et al., 2021, Kisselev et al. 2022a; Kisselev et al., 2022b; Kyle & Crossley, 2018; Kyle et al., 2020).

Human writers bring a wealth of contextual and cultural knowledge to their writings, allowing them to navigate the nuances of language that AI systems may overlook. Research has demonstrated that human-generated essays often exhibit greater syntactic complexity, as they can employ a wider variety of sentence structures and more sophisticated grammatical forms (Zindela, 2023). Translation and cross-cultural



communication abilities are now acknowledged as vital for empowering people amidst the growing global interconnection (Elmahdi & Mohamad, 2024).

Zindela (2023) highlighted the need for research that measures linguistic features to distinguish between AI-generated content and human-generated content. This study emphasizes the importance of understanding the linguistic markers that contribute to the perceived quality of writing, particularly in academic contexts. Similarly, Ha (2023) utilized computational tools to analyze syntactic complexity in written argumentation, providing insights into how different translation methods handle syntactic elaboration across various topics and language backgrounds. Using machine translation effect writing quality in addition to producing texts with linguistic features different from other texts produced by human (Ashuja'a & Jibreel, 2024; Ataman et al., 2017; Dagan et al., 2006)

In 2013, Ai and Lu adopted a corpus-based study to compare the syntactic complexity in NNS (Non Native Speakers) and NS (Native Speakers) university students' writing. The study concluded high differences in relation to the four areas of syntactic complexity; length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. The high difference in syntactic complexity between NNS students' writing and NS students' writing called for a pedagogical intervention to help the NNS students to develop their writing. Many other studies compared between syntactic complexity of texts produced by native speakers with non-native speakers (Alsahlanee & Jaganathan, 2023; Bulté & Housen, 2014)

In her study, Indarti (2018), attempted to reveal the syntactic complexity of different English editorials in different countries. Samples of texts were taken from ten online English newspapers and the samples were analyzed using the L2 Syntactic Complexity Analyzer. The native newspapers revealed the highest score of subordination which correlates with proficiency at 30 the beginning and intermediate levels. Hence, the researcher concluded that purpose of editorials was to influence the opinions of the readers on some controversial issues. Many other studies measured the syntactic complexity and its relation with L2 writing proficiency and register or genre (Deng et al., 2019; Inoue, 2016; Jagaiah et al., 2020; Larsson & Kaatari, 2020)

These studies tried to understand how syntactic complexity correlates with writing proficiency (Ortega, 2003; Norris & Ortega, 2009). They tried to distinguish between the syntactic complexity of AI-generated content and human-generated content (Zindela, 2023). Ai and Lu compared between the syntactic complexity in NNS and NS writers (2013). Indarti (2018) compared the syntactic complexity of editorials. This study explores new issues which have not been addressed. It explores the difference in syntactic complexity



between Google Translate outputs and human translations in translating texts from Arabic into English using Google Translate.

3- Methodology

Research Design

This study employs a comparative analysis to investigate the syntactic complexity of translations produced by Google Translate and human translators. The research design is structured to facilitate a systematic examination of various syntactic complexity measures across both translated texts used the two different translation methods. A web-based method approach is utilized to make quantitative analysis of syntactic features to provide a comprehensive understanding of the differences in translation quality in relation to syntactic complexity.

Participants

The study involved one professional translator and Google Translate. The human translator is selected based on their professional experience and proficiency in the source and target languages. The human translation is used as a counterpart to Google Translate to be measured against. Google Translate outputs, as a widely used machine translation tool, serve as the participant that will be measured.

Text Selection

To ensure a fair comparison, a corpus of texts is selected for translation. The texts are chosen from an Arabic book named *Hiwar Ma'a Sadeeqi Almulhid [Dialogue with An Atheist]* by Mostafa Mahmoud. Each text is approximately 100-200 words in length, providing a manageable yet substantial sample for analysis. The selected texts are translated by both the human translator and Google Translate, resulting in two sets of translations for each source text.

Syntactic Complexity Measures

The analysis of syntactic complexity is conducted using the web-based L2 Syntactic Complexity Analyzer in the website <https://aihaiyang.com/software/l2sca>. The syntactic complexity aspects include length of production unit, amount of subordination, amount of coordination and degree of phrasal sophistication. The formulas used in the website are explained in the table number one. The table shows the formulas which are used in the measuring of syntactic complexity at the different levels (Lu, 2010). These measures are selected based on their relevance to previous researches on syntactic complexity. These measures are computed as ratios indicating the level of complexity (Lu, 2011; Ortega, 2003).



Table 1

Measure	Code	Formula
<i>Length of production unit</i>		
<i>Mean length of clause</i>	MLC	# of words/# of clauses
<i>Mean length of sentence</i>	MLS	# of words/# of sentences
<i>Mean length of T-unit</i>	MLT	# of words/# of T-units
<i>Amount of subordination</i>		
<i>Dependent clauses per clause</i>	DC/C	# of dependent clauses/# of clauses
<i>Dependent clauses per T-unit</i>	DC/T	# of dependent clauses/# of T-units
<i>Amount of coordination</i>		
<i>Coordinate phrases per clause</i>	CP/C	# of coordinate phrases/# of clauses
<i>Coordinate phrases per T-unit</i>	CP/T	# of coordinate phrases/# of T-units
<i>T-units per sentence</i>	T/S	# of T-units/# of sentences
<i>Degree of phrasal sophistication</i>		
<i>Complex nominals per clause</i>	CN/C	# of complex nominal/# of clauses
<i>Complex nominals per T-unit</i>	CN/T	# of complex nominal/# of T-units

Terms in the table 1 are explained according to Lu (2010):

Sentence: A group of words starting with a capital and ends with a period, question mark, exclamation mark, quotation mark, or ellipsis.

Clause: A structure with a subject and a finite verb including adjective clauses, independent clauses, adverbial clauses, and nominal clauses.

Dependent Clauses: It is defined as a finite nominal, adjective, or adverbial clause.

T-unit: A main clause plus any non-clausal or subordinate clause structure attached to or embedded in it.

Complex T-unit: A T-unit that contains a dependent clause.

Coordinate phrase: Any noun, adverb, adjective, and verb phrases in coordinate phrases.

Complex nominals: This comprises 1- nouns plus adjectives, possessive, prepositional phrase, relative clause, participle, or appositive 2- nominal clauses 3-Gerunds and infinitives in subject position.

Verb phrase: It comprises both finite and non-finite verb phrases.

Data Analysis

The collected data is analyzed using the web-based L2 Syntactic Complexity Analyzer L2SCA (Lu, 2010) to perform a series of quantitative analyses. A quantitative analysis is conducted to show ratios indicating the



level of syntactic complexity. By comparing the ratios resulted from this computed analyzer, one can realize the differences between two texts in relation to their syntactic complexity.

4- Results

Samples for the Syntactic Complexity Analysis

Arabic Text Sample Number One

"ونحن نقول له : سؤالك فاسد .. ولا مطب ولا حاجة فأنت تسلم بأن الله خالق ثم تقول من خلقه؟! فتجعل منه خالقاً ومخلوقاً في نفس الجملة وهذا تناقض. والوجه الآخر لفساد السؤال أنك تتصور خضوع الخالق لقوانين مخلوقاته.. فالسببية قانوننا نحن أبناء الزمان والمكان. والله الذي خلق الزمان والمكان هو بالضرورة فوق الزمان والمكان ولا يصح لنا أن نتصوره مقيداً بالزمان والمكان ولا بقوانين الزمان والمكان. والله هو الذي خلق قانون السببية فلا يجوز أن نتصوره خاضعاً لقانون السببية الذي خلقه". (Mahmoud, 1986, pp. 8-9).

The Human Translation

"I replied to him by making clear that his question was meaningless. There is no dilemma or anything of that sort. You grant that God is a Creator and then you ask about who created Him making Him both creator and created in the same sentence, which is a contradiction. The other of your question's meaninglessness is that you imagine the creator as being subject to the laws which govern his creatures. Causation is a law for us who live in space and time. God, who created space and time, is necessarily transcendent in relation to both and it is an error on our part to think that he is bound either by them or by their laws. It is God who created the law of causation and we cannot consider Him as subject to the law He created". (Mahmoud, 1986/2000, pp. 5-6).

Google Translate Output

"We say to him: Your question is invalid. There is no pitfall or need for it, because you accept that God is the Creator and then you ask who created Him?! So you make Him both Creator and created in the same sentence, and this is a contradiction. The other aspect of the invalidity of the question is that you imagine that the Creator is subject to the laws of His creations. Causality is our law, we are the sons of time and place. And God who created time and place is necessarily above time and place, and it is not right for us to imagine Him bound by time and place or by the laws of time and

place. And God is the one who created the law of causality, so it is not right for us to imagine Him subject to the law of causality that He created" (Google Translate, 2024).

Table 2

Results of Sample One Analysis

Syntactic Complexity Indices	Human Translation	Google Translate
<i>Length of production unit</i>		
Mean length of clause (MLC)	7.1500	6.9091
Mean length of sentence (MLS)	20.4286	16.8889
Mean length of T-unit (MLT)	14.3000	11.6923
<i>Amount of subordination</i>		
Dependent clause per clause (DC/C)	0.6000	0.4091
Dependent clause per T-unit (DC/T)	1.2000	0.6923
<i>Amount of coordination</i>		
Coordinate phrase per clause (CP/C)	0.2000	0.2727
Coordinate phrase per T-unit (CP/T)	0.4000	0.4615
T-unit per sentence (T/S)	1.4286	1.4444
<i>Degree of phrasal sophistication</i>		
Complex nominal per clause (CN/C)	0.8500	0.6818
Complex nominal per T-unit (CN/T)	1.7000	1.1538

Arabic text sample Number Two

"وأنت بهذه السفسطة أشبه بالعرائس التي تتحرك بزمبلك. وتتصور أن الإنسان الذي صنعها لا بد هو الآخر يتحرك بزمبلك. فإذا قلنا لها بل هو يتحرك من تلقاء نفسه. قالت: مستحيل أن يتحرك شيء من تلقاء نفسه. إني أرى في عالمي كل شيء يتحرك بزمبلك. وأنت بالمثل لا تتصور أن الله موجود بذاته بدون موجد.. لمجرد أنك ترى آل شيء حولك في حاجة إلى موجد. وأنت كمن يظن أن الله محتاج إلى براشوت لينزل على البشر ومحتاج إلى أتوبيس سريع ليصل إلى أنبيائه.. سبحانه وتعالى عن هذه الأوصاف علوًّا كبيراً". (Mahmoud, 1986, p. 8).

Human Translation

"In this sophistry of yours you are like those dolls that, seeing they move by springs, imagine that the human who made them must also derive his motion from the action springs. If they were told that he is self-moved, they would retort that it is impossible for anything to move spontaneously since everything in their



world is moved by a spring. Just like them, you cannot imagine that God exists in His own Essence with no need of an efficient cause; and this is because you see everything around you in need of such a cause. It is as if you thought that God needs a parachute to descend among men or a fast car to reach His prophets; God is infinitely exalted above such conceptions". (Mahmoud, 1986/2000, p. 6)

Google Translate Output

"With this sophistry, you are like a puppet that moves with its own spring. You imagine that the man who made it must also move with its own spring. If we tell it that it moves on its own, it says: It is impossible for anything to move on its own. I see everything in my world moving with its own spring. Likewise, you do not imagine that God exists in and of itself without a creator, simply because you see that everything around you needs a creator. You are like someone who thinks that God needs a parachute to descend upon people and needs a fast bus to reach His prophets. Glory be to Him, He is far above these descriptions" (Google Translate, 2024).

Table 3

Results of Sample Two Analysis

Syntactic Complexity Indices	Human Translation	Google Translate
<i>Length of production unit</i>		
Mean length of clause (MLC)	7.7059	6.8889
Mean length of sentence (MLS)	32.7500	17.7143
Mean length of T-unit (MLT)	21.8333	15.5000
<i>Amount of subordination</i>		
Dependent clause per clause (DC/C)	0.6471	0.5556
Dependent clause per T-unit (DC/T)	1.8333	1.2500
<i>Amount of coordination</i>		
Coordinate phrase per T-unit (CP/T)	0.1667	0.1250
Coordinate phrase per clause (CP/C)	0.0588	0.0556
T-unit per sentence (T/S)	1.5000	1.1429
<i>Degree of phrasal sophistication</i>		
Complex nominal per clause (CN/C)	1.0588	0.888
Complex nominal per T-unit (CN/T)	3.0000	2.000

Arabic text sample Number Three

"أما أرسطو فقد استطرّد في تسلسل الأسباب قائلاً: إن الكرسي من الخشب والخشب من الشجرة. والشجرة من البذرة.. والبذرة من الزارع.. واضطر إلى القول بأن هذا الاستطراد المتسلسل في الزمن اللانهائي لابد أن ينتهي بنا في البدء الأول إلى سبب في غير حاجة إلى سبب. سبب أول أو محرك أول في غير حاجة إلى من يحركه. خالق في غير حاجة إلى خالق. وهو نفس ما نقوله عن الله". (Mahmoud, 1986, p. 9)

Human Translation

"Aristotle followed the chain of causal tracing the chair from wood, wood from the tree, the tree from a seed, and the seed from the planter. He had to conclude that this chain which regresses into infinite time must have begun with an 'uncaused' cause, a *primum mobile* in no need of a mover, a creator who has not been created. This is the same thing we assert of God". (Mahmoud, 1986/2000, p. 7)

Google Translate Output

"Aristotle continued in the sequence of causes, saying: The chair is made of wood, the wood is made of the tree, the tree is made of the seed, and the seed is made of the sower. He was forced to say that this sequential digression in infinite time must lead us to the first beginning with a cause that has no need for a cause. A first cause or first mover that has no need for someone to move it. A creator that has no need for a creator. This is the same thing we say about God" (Google Translate, 2024).

Table 4

Results of Sample Three Analysis

Syntactic Complexity Indices	Human Translation	Google Translate
<i>Length of production unit</i>		
Mean length of clause (MLC)	10.2857	8.3333
Mean length of sentence (MLS)	24.0000	20.0000
Mean length of T-unit (MLT)	24.0000	14.2857
<i>Amount of subordination</i>		
Dependent clause per clause (DC/C)	0.7143	0.5000
Dependent clause per T-unit (DC/T)	1.6667	0.8571
<i>Amount of coordination</i>		
Coordinate phrase per T-unit (CP/T)	0.3333	0.1429
Coordinate phrase per clause (CP/C)	0.1429	0.0833



T-unit per sentence (T/S)	1.0000	1.4000
<i>Degree of phrasal sophistication</i>		
Complex nominal per clause (CN/C)	1.5714	1.1667
Complex nominal per T-unit (CN/T)	3.6667	2.0000

Arabic text sample Number Four

"أما ابن عربي فكان رده على هذا السؤال "سؤال مَنْ خلق الخالق" .. بأنه سؤال لا يرد إلا على عقل فاسد.. فالله هو الذي يبرهن على الوجود ولا يصح أن نتخذ من الوجود برهاناً على الله.. تماماً كما نقول أن النور يبرهن على النهار.. ونعكس الآية لو قلنا إن النهار يبرهن على النور". (Mahmoud, 1986, p. 9)

Human Translation

"From another quarter, Ibn Arabi, the muslim mystic, replied to the question as to who made the creator by saying that it can only occur to a disordered mind. According to him, it is God that substantiates existence and it would be erroneous to point to existence or the universe as a proof of God. This is the same as saying that light indicates day and it would be a lopsided argument to claim that day proves the existence of light" (Mahmoud, 1986/2000, p 7).

Google Translate Output

"As for Ibn Arabi, his response to this question, "The question of who created the Creator," was that it is a question that can only be answered by a corrupt mind. God is the one who proves existence, and it is not right to take existence as proof of God. Just as we say that light proves day. We reverse the verse if we say that day proves light" (Google Translate, 2024).

Table 5

Results of Sample Four Analysis

Syntactic Complexity Indices	Human Translation	Google Translate
<i>Length of production unit</i>		
Mean length of clause (MLC)	8.3000	5.9167
Mean length of sentence (MLS)	27.6667	17.7500
Mean length of T-unit (MLT)	16.6000	17.7500
<i>Amount of subordination</i>		
Dependent clause per clause (DC/C)	0.5000	0.7500
Dependent clause per T-unit (DC/T)	1.0000	2.2500

Amount of coordination

Coordinate phrase per T-unit (CP/T)	0.2000	0
Coordinate phrase per clause (CP/C)	0.1000	0
T-unit per sentence (T/S)	1.6667	1.0000

Degree of phrasal sophistication

Complex nominal per clause (CN/C)	1.0000	0.7500
Complex nominal per T-unit (CN/T)	2.0000	2.2500

Arabic text sample Number Five

"سوف نقول له إن الخالق واحد، لأن الكون كله مبني من خامة واحدة وبخطة واحدة. فمن الأيدروجين تألفت العناصر الاثنان في فضاء الكون والتسعون التي في جدول "مندليف" بنفس الطريقة. "بالادماج" وإطلاق الطاقة الذرية التي تتأجج بها النجوم وتشتعل الشمس". (Mahmoud, 1986, p. 10).

Human Translation

"My answer was that God is one because the entire universe is built out of one material and according to a unified plan. The ninety two elements in the Mendelev table are built from hydrogen and in the same manner in which stars and suns flame-up in space; namely, by fusion and the emission of atomic energy". (Mahmoud, 1986/2000, p. 8)

Google Translate Output

"We will tell him that the Creator is one, because the entire universe is built from one material and with one plan. From hydrogen the two elements in the universe and the ninety in Mendeleev's table were formed in the same way. "by merging" and releasing the atomic energy that fuels the stars and ignites the suns" (Google Translate, 2024).

Table 6

Results of Sample Five Analysis

Syntactic Complexity Indices	Human Translation	Google Translate
<i>Length of production unit</i>		
Mean length of clause (MLC)	15.0000	8.5714
Mean length of sentence (MLS)	30.0000	12.0000
Mean length of T-unit (MLT)	30.0000	8.5714
<i>Amount of subordination</i>		



Dependent clause per clause (DC/C)	0.5000	0.2857
Dependent clause per T-unit (DC/T)	1.0000	0.2857
<i>Amount of coordination</i>		
Coordinate phrase per T-unit (CP/T)	1.0000	0.1429
Coordinate phrase per clause (CP/C)	0.5000	0.1429
T-unit per sentence (T/S)	1.0000	1.4000
<i>Degree of phrasal sophistication</i>		
Complex nominal per clause (CN/C)	1.5000	1.0000
Complex nominal per T-unit (CN/T)	3.0000	1.0000

5- Discussion

In relation to length of production unit level, Google Translate results in lower syntactic complexity ratios than human translation in all five samples. For the amount of subordination level, Google Translate results in lower syntactic complexity in all samples except in sample number four. For the amount of coordination level, Google Translate results in lower syntactic complexity in most samples with exception in sample one. However, Google Translate results in higher complexity in relation to the T-unit per sentence in samples number one and three. In relation to the degree of phrasal sophistication, Google translate results in lower complexity ratios in all samples except in complex nominal per T-unit measure in sample number five.

6- Conclusion and Implications

This study has provided comprehensive comparison of syntactic complexity in texts translated by Google Translate and human translation in the following four areas: length of production unit, amount of subordination, amount of coordination and degree of phrasal sophistication. The results indicated statistically significant differences in syntactic complexity between Google Translate and human translation in most of the 10 measures analyzed, with few exceptions as the results showed. The results of the study reveal that Google Translate output has less syntactic complexity than human translation and these findings provide an answer to the first question that the paper attempted to answer.

The findings of this research will likely to evaluate Google Translate performance in accurately handling complex syntactic structures. This can inform the future development of machine translation systems, highlighting the need for improved syntactic processing capabilities and context-awareness in MT tools. It may push developers to refine algorithms for better syntactic complexity handling.



By comparing the syntactic complexities of machine and human translations, the research could provide valuable insights for Google Translate users. This can make them able to distinguish the translations produced by Google Translate from others produced by human. By recognizing the syntactic complexity features of Google Translate output, the user can distinguish between Google Translate translation and human translation. This provide an answer to the second question in the research.

7- Recommendations for Future Research and Limitations

The research suggests conducting researches on hybrid translation approaches combining human expertise with MT tools. A hybrid translation model, where machines handle the basic structure and speed of translation, while human translators refine the output for accuracy and fluency, could be an optimal solution. This model might be especially useful for businesses and industries needing to process large volumes of text quickly while still ensuring high-quality output. Future research also should explore and analyze the accuracy of audio translation applications which are important for developing international communication.

The study may be limited to the language pair of Arabic and English. The direction of translation is limited to Arabic to English translation. The scope of syntactic complexity is limited to length of production unit, amount of subordination, amount of coordination and degree of phrasal sophistication.

References

- Ai, H., & Lu, X. (2013). *A Corpus-based Comparison of Syntactic Complexity in NNS and NS University Students' Writing*. Amsterdam: John Benjamins.
- Alnasery, I. H. S. (2024). Reducing the Ambiguity in Translating Prepositions from English into Arabic. *Arts for Linguistic & Literary Studies*, 6(2), 541–555. <https://doi.org/10.53286/arts.v6i2.1959>
- Alsahlane, A., & Jaganathan, P. (2023). Phrasal Syntactic Complexity Measures in Linguistics Research Articles Written by Iraqi and English L1 Writers. *3L the Southeast Asian Journal of English Language Studies*, 29(1), 154–168. <https://doi.org/10.17576/3L-2023-2901-11>.
- Ashuja'a, A., & Jibreel, I. (2024). Translator Praxis: An Investigation into the Practical Component in BA Translation Programs at Yemeni Universities. *Arts for Linguistic & Literary Studies*, 6(3), 574–604. <https://doi.org/10.53286/arts.v6i3.2085>
- Ataman, D., Negri, M., Turchi, M., & Federico, M. (2017). Linguistically Motivated Vocabulary Reduction for Neural Machine Translation from Turkish to English. *the Prague Bulletin of Mathematical Linguistics*, 108(1), 331–342. <https://doi.org/10.1515/pralin-2017-0031>
- Bulté, B., & Housen, A. (2014). "Conceptualizing and measuring short-term changes in L2 writing complexity." *Journal of Second Language Writing* 26, 42-65.



- Cancino, M., & Panes, J. (2021). The impact of Google Translate on L2 writing quality measures: Evidence from Chilean EFL high school learners. *System*, 98, 102464. <https://doi.org/10.1016/j.system.2021.102464>
- Dagan, I., Glickman, O., & Magnini, B. (2006). The PASCAL Recognising Textual Entailment Challenge. In *Lecture notes in computer science* (pp. 177–190). https://doi.org/10.1007/11736790_9
- Deng, Y., Lei, L., & Liu, D. (2019). Calling for More Consistency, Refinement, and Critical Consideration in the Use of Syntactic Complexity Measures for Writing. *Applied Linguistics*, 42(5), 1021–1028. <https://doi.org/10.1093/applin/amz069>
- Elmahdi, O. E. H. & Mohamad, H. M. H. (2024). Developing Translation and Interpretation Skills: An Analysis of Students' Linguistic Needs, *Arts for Linguistic & Literary Studies*, 6(4): 652–671
- Google Translate. (2024, October 1). In Wikipedia. https://en.m.wikipedia.org/wiki/Google_translate.
- Ha, M.J. (2023). Computer-Aided Analysis of Syntactic Elaboration of Written Argumentation Across Topics and L1s. *Computer Assisted Language Learning Electronic Journal CALL-EJ*, 24(1), 242–268
- Hwang, H., Jung, H., & Kim, H. (2020). Effects of Written Versus Spoken Production Modalities on Syntactic Complexity Measures in Beginning-Level Child EFL Learners. *Modern Language Journal*, 104(1), 267–283. <https://doi.org/10.1111/modl.12626>
- Inoue, C. (2016). A comparative study of the variables used to measure syntactic complexity and accuracy in task-based research. *Language Learning Journal*, 44(4), 487–505. <https://doi.org/10.1080/09571736.2015.1130079>
- Jagaiah, T., Olinghouse, N. G., & Kearns, D. M. (2020). Syntactic complexity measures: variation by genre, grade-level, students' writing abilities, and writing quality. *Reading and Writing*, 33(10), 2577–2638. <https://doi.org/10.1007/s11145-020-10057-x>
- Kisselev, O., Klimov, A., & Kopotev, M. (2021). Syntactic Complexity Measures as Indices of Language Proficiency in Writing: Focus on Heritage Learners of Russian. *Heritage Language Journal*, 18(3), 1–30. <https://doi.org/10.1163/15507076-12340016>
- Kisselev, O., Klimov, A., & Kopotev, M. (2022a). Syntactic complexity measures as linguistic correlates of proficiency level in learner Russian. In *Studies in corpus linguistics* (pp. 51–80). <https://doi.org/10.1075/scl.104.03kis>
- Kisselev, O., Klimov, A., & Kopotev, M. (2022b). Syntactic complexity measures as linguistic correlates of proficiency level in learner Russian. In *Studies in corpus linguistics* (pp. 51–80). <https://doi.org/10.1075/scl.104.03kis>
- Kyle, K., & Crossley, S. A. (2018). Measuring Syntactic Complexity in L2 Writing Using Fine-Grained Clausal and Phrasal Indices. *Modern Language Journal*, 102(2), 333–349. <https://doi.org/10.1111/modl.12468>
- Kyle, K., Crossley, S., & Verspoor, M. (2020). MEASURING LONGITUDINAL WRITING DEVELOPMENT USING INDICES OF SYNTACTIC COMPLEXITY AND SOPHISTICATION. *Studies in Second Language Acquisition*, 43(4), 781–812. <https://doi.org/10.1017/s0272263120000546>
- Larsson, T. & Kaatari, H. (2020). Syntactic complexity across registers: Investigating (in)formality in second-language writing. *Journal of English for Academic Purposes* 45, 1–16 <https://doi.org/10.1016/j.jeap.2020.100850>



- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4): 474-496.
- Lu, X. (2011). A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *TESOL Quarterly*, 45(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, 34(4), 493–511. <https://doi.org/10.1177/0265532217710675>
- Machine translation. (2024, May 18). In Wikipedia. https://en.m.wikipedia.org/wiki/Machine_translation.
- Mahmoud, M. (1986). *Hiwar Ma'a Sadeeqi Amulhid* [Dialogue With An Atheist]. Dar Ala'awda, Beirut.
- Mahmoud, M. (2000). *Dialogue With An Atheist* (M, Yehia, Trans.). Dar Al Taqwa. (1986).
- Norris, J.M. and Ortega, L. (2009), "Measurement for understanding: An organic approach to investigating complexity, accuracy, and fluency in SLA", *Applied Linguistics*, 30, 555–578.
- Ortega, L. (2003). Syntactic Complexity Measures and their Relationship to L2 Proficiency: A Research Synthesis of College-level L2 Writing. *Applied Linguistics*, 24(4), 492–518. <https://doi.org/10.1093/applin/24.4.492>
- Park, S. (2012). A Corpus-Based Study of Syntactic Complexity Measures as Developmental Indices of College-Level L2 Learners' Proficiency in Writing. *Korean Journal of Applied Linguistics*, 28(3), 139. <https://doi.org/10.17154/kjal.2012.09.28.3.139>
- Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, 28, 53–67. <https://doi.org/10.1016/j.jslw.2015.02.002>
- Yu, Q. (2019). An Organic Syntactic Complexity Measure for the Chinese Language: The TC-Unit. *Applied Linguistics*, 42(1), 60–92. <https://doi.org/10.1093/applin/amz064>
- Zeng, X. (2023). Comparing Syntactic Complexity: Research Papers Written by Undergraduate Researchers and Published Research Papers. *Journal of Education*, 203(1) 129-138. <https://doi.org/10.1177/00220574211016435>
- Zindela, N. (2023). Comparing Measures of Syntactic and Lexical Complexity in Artificial Intelligence and L2 Human-Generated Argumentative Essays. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)* 19(3), 50-68.

