



A Computational Analysis of Syntactic Simplification in Gen-Z Digital Communication Using NLP Tools

Javera Ahmed¹, Nisar Ahmed², Mifra Hussain³

¹MS Scholar, COMSATS University, Islamabad – Lahore Campus. javeahmed16@gmail.com

²MS Scholar, COMSATS University, Islamabad – Lahore Campus. baladinisar18@gmail.com

³MS Scholar, COMSATS University, Islamabad – Lahore Campus. hussainmifra2@gmail.com

ARTICLE INFO

Article History:

Received: August 06, 2026

Revised: August 13, 2026

Accepted: August 27, 2026

Available Online:

Keywords:

Gen Z communication, syntactic simplification, computational linguistics, NLP analysis, cognitive load theory

Corresponding Author:

Javera Ahmed

Email: javeahmed16@gmail.com

ISSN Print: 3005 - 5644

ISSN Online: 3005 - 5652

Volume 4, No. 1, 2026, Pages 61 – 83

Home Page: <https://journals.sbbusba.edu.pk>

This is an open access article under the CC BY-SA license.



Copyright © 2024 by author, Published by Shaheed Benazir Bhutto University, Shaheed Benazirabad.

ABSTRACT

In this paper, syntactic simplification patterns in Generation Z digital communication are analyzed based on the computational linguistic analysis of the data with Natural Language Processing (NLP) tools. The data set used in this research is a set of 160 comments obtained in one of the discussion threads in the r/Law School sub-community on the Reddit platform, which is an informal online communication mainly between students of universities and young adults. The corpus that was gathered consists of 10,476 words and was processed in Python with NLP libraries, such as NLTK, spacey, Stanza, and Text Blob. These linguistic aspects analyzed were the sentence length, the part-of-speech allocation and frequency of words in order to determine the patterns of simplified syntactic constructions. The findings indicate that Gen Z digital communication often depends on shorter sentence constructions, extensive use of pronouns and verbs, and frequent use of function words that help to maintain grammatical coherence and limit the use of complex lexical meanings. These tendencies are then explained within the framework of the Cognitive Load Theory that explains that the restrictions of working memory prompt speakers and writers to simplify language in communication situations where the speed is important. The results indicate that syntactic simplification in online communication is not linguistic loss but a survival technique due to cognitive and communicative resource limitations.

INTRODUCTION

In the modern digital world, the communication of younger generations shifts towards a more online messaging platform, namely, WhatsApp, Instagram, and the other social media apps. Such sites have played a great role in defining the process of language production and consumption in everyday communication. Users of the Gen Z generation tend to depend on brief and fast communication in the digital messages, which can often lead to a simpler structure of sentences, exclusion of grammatical features, and fragmented phrases. Linguists have observed that the digital communication contexts promote more lax and cost-effective language usage than the traditional written way of communication (Crystal, 2008).

The study of language in digital form usually referred to as internet linguistics or computer-mediated communication (CMC) has evolved through the last several decades. It started to be researched in the 1980s and 1990s when first text-based platforms, such as Usenet, email, and bulletin boards, became studied, and scholars investigated how users modified language to fit technological limitations like speed and message length (Panajaitan & Patria, 2024). The emergence of mobile devices and social media in the 2000s and social media in the 2010s changed research to focus on patterns of online language use, including the use of abbreviations, slang, and simplified sentence structures (Stancescu, 2025). Furthermore, more recent researchers note that digital communication still affects the sentence construction, the brevity, and the informality, and the study of the syntactic patterns in the online messages represents a significant line of investigation (Azaz et al., 2024).

Syntactic simplification has been popularly debated in both computational linguistics and pedagogical research as the act of diminishing the complexity of a text through its grammatical structure and the preservation of its original meaning and information structure. It is usually connected with the simplification of the sentence structures that are complex in their nature (they may be subordinated, coordinated, contain relative clauses, appositions, or are passive voice-structured, etc.) into shorter and simpler sentences (Ivchenko, 2022, p. 634). Computationally, syntactic simplification was first suggested by Chandrasekar et al. (1996) as a preprocessing step to automated parsers to make the sentences shorter in order to achieve a higher parsing rate. Siddharthan (2006) went further by adding discourse-sensitive conveyances, permuting simplified

sentences, choosing rightful determiners, and introducing rhetorical linking to make sure of a coherent output, and retained the output to a broader audience. All these studies demonstrate that syntactic simplification is an important process of not only improving the comprehension process but also complementing the computational processing of language.

The use of Natural Language Processing (NLP) is specialized software, libraries, and frameworks that allow computers to read, comprehend, create and manipulate human language. NLP is a subdiscipline of computer science and artificial intelligence (AI) that integrates computational linguistics, rule-based language modeling, and uses statistical and deep learning to process text and speech (Cole and Holdsworth, n.d.; Awati, Gillis and Lutkevich, 2025). NLP can be traced to the 1950s when shortly after World War II the initial attention was made to provide machines with the ability to translate languages and to comprehend human text (Polzin, 2022). The systems of the 1960s and 1970s, ELIZA and SHRDLU, were rule-based and depended on manually written rules of grammar to simulate human language knowledge (GeeksforGeeks, 2025). As statistical methods grew in expertise by the 1980s and 1990s, purely rule-based methods began to be gradually supplanted by machine learning algorithms which trained patterns on large corpora of data, allowing more scalable and effective methods to be used in speech recognition and machine translation (Box, 2025).

The Generation Z, in general, being born between the late 1990s-early 2010s, is the first generation to grow up in a fully digitalized space and social media, like Instagram (Telaumbanua et al., 2024). Being always connected and digitally fluent they encourage unique language practices, which are typified by the high frequency of the use of abbreviations, slang, and new spelling patterns, and multimodal forms of communication, which involves an integration of text and images (Marzona, 2025). These are sometimes called netspeak or leetspeak, in which letters are substituted with numbers or symbols to represent meaning in a short amount of time or evade content-moderating algorithms (Beck, 2025). Although these innovations can make the process of immediacy and identity expression only online possible, researchers have expressed concerns that these practices contribute to the undermining of language, rendering it more simplistic and semantically less robust, thereby making communication less straightforward and predictable and more representative of the wider Gen Z tendencies towards language negotiation within the online space (Javeed, 2025).

Cognitive Load Theory (CLT) is a concept that was created by John Sweller in the late 1980s and has a theoretical foundation of human working memory having a limited capacity to process information at any one given time. According to the theory, cognition resources become overloaded when people are provided with excessive information or complicated systems, and this can influence the learning and comprehension. To cope with this shortcoming, CLT focuses on structuring information in a way that minimizes the mental effort that people do not need and permits efficient processing of information. In this regard, the theory emphasizes how human beings tend to favor easier and less demanding systems of information processing, especially within a busy communication setup (Clark & Kimmons, 2023).

This research falls under the field of syntax in computational and applied linguistics, which involves the simplification of sentence structures in the digital communication of Generation Z. Gen Z are often used to send messages that are rather short, fast, and consist of text messages. These contexts usually promote shortened forms, smaller sentence structures and broken constructions which are not in accord with conventional syntactic patterns. These developments are also causing concern on how the digital communication is affecting the structure of the sentence and the complexity of language. The proposed research is aimed at examining the trends in syntactic simplification in Gen Z digital communication with the help of Natural Language Processing (NLP) tools and assessing the impact of digital messaging contexts on the organization of the sentences of young generations.

OBJECTIVES

To reach this goal, the following research objectives are suggested:

- To identify patterns of syntactic simplification in Generation Z digital messages.
- To analyze syntactic structures in digital communication using NLP tools.
- To explore the relationship between syntactic simplification and cognitive effort in digital messaging.

RESEARCH QUESTIONS

The present study addresses the following research questions:

1. What types of syntactic simplification are commonly found in Generation Z digital communication?
2. How can NLP tools be used to analyze syntactic structures in Gen Z digital messages?
3. What relationship exists between syntactic simplification and cognitive effort in digital messaging?

LITERATURE REVIEW

The issue of how digital communication affects language change is a topic that has received much debate in linguistic studies. Crystal (2008) observes that online communication tends to encourage shorter sentences, abbreviations and unstandardized grammar because of the speed and immediacy of digital communication. Such tendencies in linguistics indicate that online communication can tend to affect the syntactic frames of the older generations, as time goes by. A number of researchers have particularly studied the linguistic behavior of Generation Z on the internet. Telaumbanua et al. (2024) observe that their online communication is often full of abbreviations, slang, and inventive spelling due to the fact that Generation Z is raised in a highly connected digital space. These attributes are commonly linked to the aspects of netspeak, which is a type of language that is characterized by speed and efficiency.

On the same note, Marzona (2025) discovered that the vocabulary and language use amongst Gen Z learners was highly influenced by the digital platform, where language use is generally short and informal. These results show that online spaces promote language creativity and make the linguistic system simpler at the same time. The use of digital communication in promoting structural simplification in sentences has been pointed out by other researchers. A study carried out by Tagliamonte and Denis (2008) examined conversations via instant messaging and determined that online communication has a high number of individuals who minimize word and phrase structures and use shorter sentences. Their analysis is that linguistic economy, where simpler structures are used in communicating more effectively, is encouraged by the immediacy of digital communication, in which speakers use simpler structures. These observations lend credence to the view that online communication spaces could have a role to play in the complexity of syntactic structures that are applied in the production of everyday language.

Jabeen and Nawaz (2025) studied the morphological innovations in the language used by Generation Z through the analysis of the linguistic pattern placed in social media, twitters, Tik Tok, and Instagram. The research based on a small corpus study and ethnographic interviews with Gen Z, Pakistani-based participants discovered multiple word-formation processes such as coinage, borrowing, compounding, blending, clipping, and acronyms. Other mechanisms including contractions, reduplications and orthographic changes were also found in the findings. The scholars came to the conclusion that spelling changes and imaginative word building are widespread among the users of Gen Z, which is an indication of the dynamic English vocabulary in the realm of the digital communication process.

The second study was done by Lorenzo (2025) who investigated the connection between knowledge of Generation Z vocabulary and English language proficiency among Grade 11 students. The study conducted on the familiarity of Gen Z vocabulary among 151 participants with a descriptive-correlational type of research established a strong positive correlation between the familiarity of Gen Z vocabulary and the levels of English language proficiency among the students. The findings indicated that students who are more conversant with modern slang and computer vocabularies are likely to have better language proficiency. These results show that the current tendencies in the field of language applied by younger generations can have an impact on language development and learning.

In a similar study, Tufail et al. (2024) did a study to understand the popularity and usage of slang words by users in the generation Z social media. Through qualitative content assessment, the research examined the trends of communication among university students and discovered that slang is commonly used to establish intimacy and enhance social interaction during online communication. The authors found a variety of forms and communicative purposes of slang and suggested that technological advancement and social media sites contribute significantly to the formation of linguistic activities of Gen Z users.

Studies have also been done on how language use is related to the aspect of generational identity. Yeasmin (2025) also evaluated the role of generation specific jargon in the development of a generational identity among bilingual speakers. Through a mixed approach analysis, the authors discovered that the application of Gen Z specific expressions and language patterns can be

used as an indicator of group identity and can be used to distinguish between younger speakers and other generations. The results demonstrate that language is not only a means of communication, but it also shows a social and generational identity.

Sitohang and Ristia (2025) offered additional information on the Gen Z language practices by examining semantic shifts in slang vocabulary used by Gen Z in social media platforms, including Tik Tok, Instagram, and Twitter. In the qualitative analysis and ethnographic interviews, Gen Z slang often experiences semantic change including pejoration, amelioration, metaphorical extension, and semantic narrowing, which the research discovered. They were discovered to have a number of communicative values in these linguistic innovations such as humor, emotional authoring, constructed identities, and social bonding in digital spaces.

In a similar manner, Telaumbanua et al. (2024) examined the language manner that Generation Z teenagers use when posting on Instagram. According to their findings, the language style of Gen Z communication is mostly informal in nature with the use of language forms such as slang, abbreviations, and colloquial expressions. The paper also established various forces in this linguistic pattern and they included technological affordances, social interaction, and the changing trends in language among the digital communities. These results indicate that social media sites have a considerable influence on language behavior among the young generations.

Gen Z brevity of communication has also been mentioned in media discussions. A report released in Dawn (2025) pointed out that one of the most prominent features of Gen Z lingo is its conciseness, which was partly a byproduct of the initial social media that restricted the number of words to use. Such limitations promoted the adoption of abbreviations and reduced phrases like "LOL," which led to the development of a communication culture based on speed and efficiency.

Despite the extensive literature on the topic of vocabulary, slang and morphological novelties used by the Generation Z in social media as the means of communication, little is known about the syntactic patterns and simplification of the sentences used by the cohort. Although the occurrence of abbreviations, informal language, and lexical innovations has been investigated, very frequently the authors do not mention how and why Gen Z simplifies the sentence structure in online communication. It is important to note that this gap is especially important since syntactic

simplification can indicate underlying cognition, including the reduction of cognitive load in the context of high-speed messaging, but current research has not been computationally manipulated with NLP tools. Thus, the following paper intends to bridge this gap by exploring the syntactic simplification techniques used by the Generation Z in online communication and to discuss them through the prism of the Cognitive Load Theory to determine why such simplifications take place in the first place.

RESEARCH METHODOLOGY

Theoretical Background

The present research paper uses Cognitive Load Theory (Sweller, 1988) as the primary theoretical framework. This theory describes the role of constraints in the human cognitive processing on the manner in which individuals interpret and generate information. The Cognitive Load Theory postulates that working memory does not have much processing capacity of information at a given time and when that capacity is surpassed people become affected by cognitive overload. Consequently, individuals tend to engage in methods that help them to lower the mental demand and simplify information processing (Clark and Kimmons, 2023). It can be applied to explain the reasons why people could be drawn to a simpler linguistic form especially in a hasty communication setting.

In the framework of the digital communication Cognitive Load Theory can be used to understand why the Generation Z users use shorter phrases, skip some grammatical forms and use simplified syntactic frameworks in the digital communication. Since communicating online tends to be faster and information processing is quicker, the structures of the sentences used by the speakers can be simplified in order to utilize less cognitive input and achieve more productive communication. By means of this, Cognitive Load Theory also offers an insightful way of understanding the syntactic simplification trends in Generation Z digital communication that are subsequently computed with the help of NLP tools in the given work.

Research Design

A mixed-method research design, comprising of computational quantitative analysis and theoretical interpretation, is applied in the research and used to investigate the tendencies of syntactic simplification in Generation Z digital communication. The research uses Natural Language Processing (NLP) applications to make a quantitative analysis of language structures and the findings are explained by the theoretical frame of Cognitive Load Theory (Sweller, 1988). This theoretical approach describes the role of constraints in human working memory in language processing and promotes the adoption of simplified structures in high-paced communication settings.

The data to be analyzed is online comments of one of the discussion threads on the Reddit platform that is popular with university students and young adults. The data was obtained on the basis of a reddish post of the user Additional-Reserve66 on the r/Law School community (Additional-Reserve66, 2025). The chosen post, which is entitled I love watching my test anxiety friends walk out of an exam 2 hours after I had to be, was published about three months before the data collection process, and includes discussion comments related to the experience of exams, in specific. One hundred and sixty-one comments were codified into a text corpus. The preprocessed corpus had 10,476 words and an average of 65.47 words per comment.

The chosen discussion thread is a natural conversation between university students and young adults, which makes it possible to observe the use of informal language and the syntactic patterns used by the Generation Z in the online dialogue. It is not the methodology that tries to check the generational identity of every commenter, instead, one would use the contextual appropriateness of the thread to a student-oriented community that can be somewhat attributed to the Generation Z group. The obtained corpus was processed with the help of Python and a few NLP packages, such as NLTK, spacey, Stanza, and Text Blob. The tools were applied to analyze the syntactic characteristics, including the length of the sentences, the structure of tokens, and the distribution of part-of-speech to determine the trends of simplified constructions in the samples. The computational results were then discussed in terms of Cognitive Load Theory, which will offer a theoretical explanation of why streamlined sentence structure can be used in quick-paced digital communication.

DATASET AND SAMPLING

The sample size consists of 160 comments of the chosen Reddit thread. This thread was chosen with the help of purposive sampling since it centers on the discussion pertinent to the experiences of exams as a student group, which is why the content is created by members of the group that is often discussed as a part of Generation Z. This will enable the study to concentrate on informal written communication among young adults in a student setting, which has a direct relation to the research objectives. The comments were all open and there was no actual personal communication with the users. Preprocessing was done to remove the identifiable information, including usernames, to keep participants anonymous. The study focused on a newly created post (within three months) and this guarantee the data obtained is a current communication practice among Generation Z students.

DATA ANALYSIS

The comments obtained on Reddit underwent computational linguistic examination. The first step involved in the process was the preprocessing of the text corpus through the elimination of empty lines and irrelevant materials, which was then followed by the use of NLTK in tokenizing words and sentences and basic text normalization to make the text readable by the syntactic analyzer. Following preprocessing, the corpus was processed with NLTK, spacey, Stanza and Text Blob in order to investigate the essential linguistic aspects such as sentence structure, token patterns and part-of-speech distributions. This discussion aimed at defining the trends of syntactic simplification that is widespread in informative and rapid digital communication between young adults. Findings of these analyses are discussed in the framework of Cognitive Load Theory (Sweller, 1988) that indicates that the cognitive constraints can stimulate shorter and simpler sentence constructions in online messaging. This model can be used to justify the fact that Generation Z consumers can simplify their syntax in non-formal digital communication.

ETHICAL CONSIDERATION

Any comments that were used were publicly posted and no individual was approached to participate. Information that could be identified was eliminated in order to preserve anonymity. The data is collected only to conduct academic research and no information is provided to any

third party. There were also ethical rules to online text data utilization, which made the study qualify as standard research in terms of ethics.

FINDINGS

Overview of the Dataset

The data used in this research comprises of 160 Reddit posts in one of the discussion threads focused on exams and dedicated to students. Prior to examination, all the extraneous spaces and irrelevant data were eliminated so that the textual information that has been put across clearly expressed the language use among Generation Z respondents.

Table 4.1: Dataset consisted of 10,476 words and the average number of words per comment is 65.47

Category	Value
Total Comments	160
Total Words	10,476
Average Words per Comment	65.47

As can be seen in Table 4.1, the comments gathered by them have enough text to perform syntactic analysis. The average number of words per comment is rather high, which means that, although these messages are digital and informal, Gen Z users can create full-blown sentences in which syntactic patterns, such as simplifications, omissions, or simplified clause constructions, can be observed.

Sentence Length Analysis

Sentence length was to be analyzed to get a clear picture of the complexity of structures in data syntactically.

Table 4.2: Marked off into sentences and the quantity of words per sentence

Category	Value
Total Sentences	450
Average Sentence Length	23.28 words
Shortest Sentence	2 words
Longest Sentence	119 words

As indicated in table 4.2, the mean sentence length of 23.28 words implies the shortest sentence of 2 words and longest sentence of 119 words. These findings indicate that digital messages tend to differ greatly in terms of syntactic complexity. Short sentences, which are prevalent, can be evidence of intentional simplifications or omissions that are often found in informal online communication, which is consistent with previous literature indicating that the Gen Z likes speed and cognitive convenience over complete syntactic structures. Sentences that are longer, however, occur when the users are describing an experience or explaining something in detail.

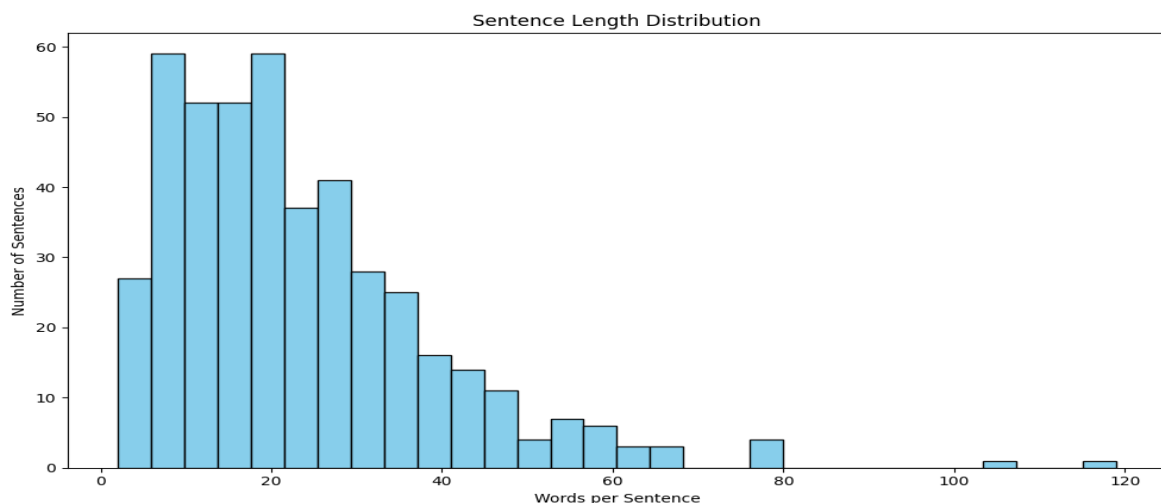


Figure 4.1: Simplification of syntax is not always the same when applied in all messages.

Part-of-Speech Distribution

The analysis of the part-of-speech was performed to reveal the patterns of the syntactic use and possible simplifications.

Table 4.3: The patterns of the syntactic use and possible simplifications.

POS Tag	Frequency
PRON	1242
VERB	1392
AUX	723
ADP	873
CCONJ	335
PART	380
ADJ	724

DET	741
NOUN	1912
PUNCT	935
SCONJ	313
NUM	77
INTJ	35
ADV	597
PROPN	87
X	14
SYM	13
SPACE	1

As it is seen in Table 4.3, nouns (1,912) and verbs (1,392) prevail in the dataset, in line with the messages that revolve around subjects and actions. The number of occasions on which pronouns are used is high (1,242), which shows that the comments are personal and experiential. The relatively smaller number of auxiliaries (723), conjunctions (335), and prepositions (873) indicate a possibility that a lot of the sentences leave out some grammatical features, which can prove the hypothesis of simplifying the syntax of digital messages. This tendency is in line with the fact that Gen Z tends to write structurally simpler, even though semantically complete sentences, and this tendency indicates a sign of lower mental expenditures in sentence construction.

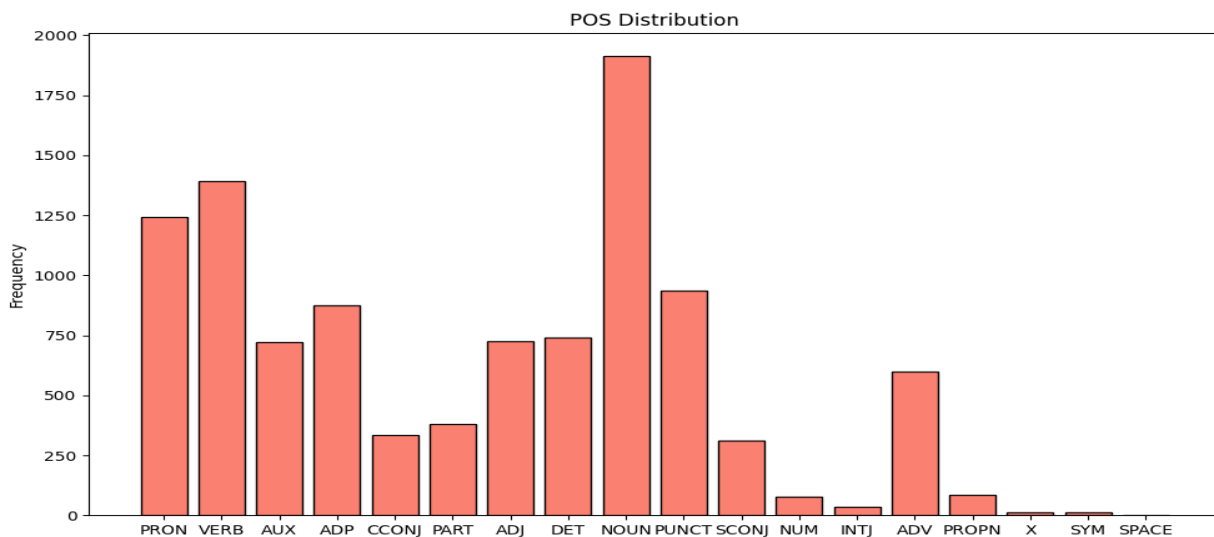


Figure 4.2: tendency indicates a sign of lower mental expenditures in sentence construction.

Most Frequent Words

The frequency of the words was assumed to find out the most widespread words and how those could be connected with the syntactic patterns (Table 4.4).

Table 4.4: The frequency of the words and how those could be connected with the syntactic patterns.

Word	Frequency
The	309
To	254
A	199
I	193
And	188
That	177
Of	155
You	145
It	140
Is	127

Analysis demonstrates that the most common words are those functional in nature, including the, to, a, and the personal pronouns, including I and conjunction, including and. These findings suggest that messages often have fundamental structural features required in meaning but often not optional or more complicated syntactic structures. The widespread use of personal pronouns and straightforward verbs also indicate the conversational and experience nature of Gen digital communication as these forms is focused on clarity and concision rather than embedding complex sentences.

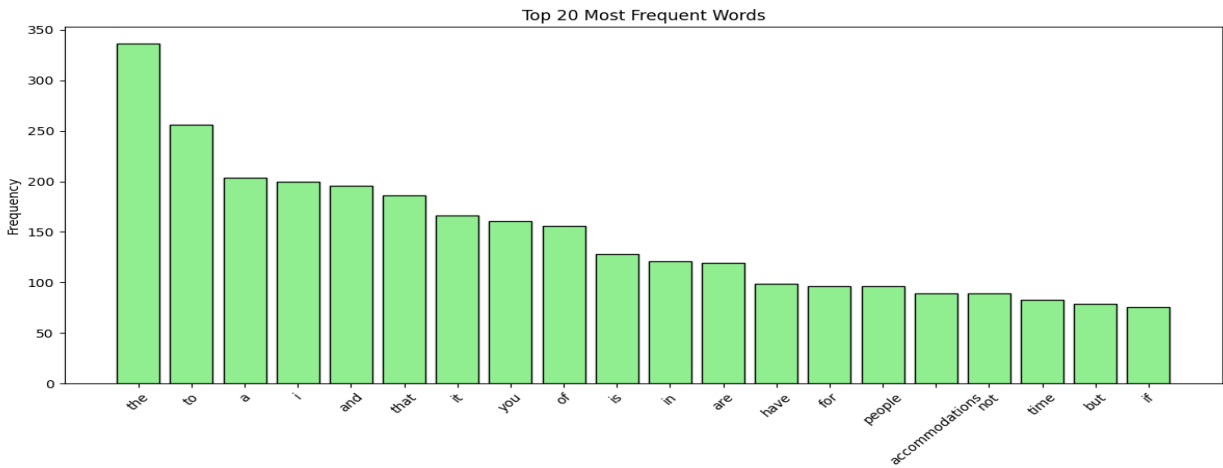


Figure 4.3: Nature of Gen digital communication as these forms is focused on clarity

DISCUSSION

The syntactic simplification of modern online communication is also interestingly analyzed by the analysis of 160 Gen Z digital comments (10,476 words). The mean sentence length of 23.28 words and a range of 2 to 119 words indicate that though some of the messages are rather brief, some of them nevertheless include complicated structures. Such variation is in line with the projections of Cognitive Load Theory (Sweller, 1988; de Jong, 2010) which assume that human working memory is limited in capacity and that people naturally develop strategies to help them lower cognitive load in real time language processing and production. The fact that the sentences are shorter and have simpler syntactic structure can be explained in the context of digital messaging by the need to control intrinsic and extrinsic cognitive load. The fact that pronouns (PRON = 1242), verbs (VERB = 1392), and auxiliary verbs (AUX = 723) dominate the prevalence of other parts of speech, i.e. adjectives (ADJ = 724) and proper nouns (PROPN = 87), is indicative of the interest in communicating only the necessary information. This is consistent with the idea that in situations of limited cognitive resources users first focus on the basic components that they need to make meaning such that they end up minimizing on the unnecessary cognitive work (Clark and Kimmons, 2023; Sweller, 2011).

The large proportion of the frequentness of function words, like the, to, a, I, and (309, 254, 199, 188, respectively) also suggests that people tend to create sentences that are grammatically coherent as well as unsaturated in terms of content. This can be explained by Cognitive Load Theory as a measure to cope with intrinsic cognitive load, which is set by the complexity of information being processed by its nature (Sweller, 2011). In online messaging, the message content is usually contextually communicated and known, hence the working memory is not required to process new information. The dependency on the words that are very simple and frequent indicates that Gen Z users are strategically distributing their cognitive resources to convey their main ideas without adding too much load to their working memory. This observation is complemented by de Jong (2010) who highlights working memory capacity limitations to the shaping of information by human beings especially in real-time communication situations.

Those patterns of simplification in the syntactic construction are further emphasized by the POS distribution. A greater usage of verbs and pronouns is also associated with an emphasis on action and participants whereas the relatively less usage of conjunction (CCONJ = 335) and subordinating conjunction (SCONJ = 313) is related to a lack of the complex syntactic structures like compound and embedded clauses. The Cognitive Load Theory is based on the idea that the processing of such complex structures elevates extraneous cognitive load, which does not have anything to do with learning or the understanding of messages (Sweller, 1988; Clark and Kimmons, 2023). The reduction of cognitive load on the writer of the message and the recipient of the communication is well achieved through message writers through the sparse use of conjunctions and subordinating clauses, which favors the concept that digital communication is more focused on message processing ease rather than syntactic richness.

Also, the statistics reveal selective usage of adjectives (ADJ = 724) and adverbs (ADV = 597) that are commonly incorporated to state or qualify something. Although these elements make communication more interesting and diverse, their small share in comparison with verbs and nouns indicates that Gen Z communicators balance the expression of detail and cognitive effectiveness strategically. Such selective involvement corresponds to principles of germane cognitive load which Sweller defines as the effort applied to the construction and automation of schemas in long-term memory (de Jong, 2010). In electronic messaging, the "schemas" could be understood as the common knowledge or cultural awareness between the sender and the receiver. Using the key action-oriented words and key references, the user maximizes the functionality of working memory, so that the mental effort that is spent is concentrated on meaningful communication but not on non-essential processing.

This economy of cognitive effort can be easily seen by the shortest statements of the dataset (some are only two words long). These ultra-coded messages are probably replies in a conversation thread that is in progress, and where background is already known and there is little working memory pressure. On the other hand, the maximum sentence of 119 words shows that the cognitive load is not always the same and that sometimes, users will construct more elaborately the sentence when the content requires more intrinsic cognitive presence or when the sender would like to be more detailed. According to Sweller (1998), intrinsic cognitive load is affected by the interactivity and complexity of factors in a task. In messaging, more interactively oriented or multi-dimensional

concepts might take longer to construct in a way that is understandable, even though this can be cognitively expensive. Such pattern confirms that simplification is relative and depends on the context, and this is an adaptive process of dealing with cognitive load.

The analysis of the most common words in conjunction with the POS distribution indicates that Gen Z is focused not on the lexical richness but on communicative efficiency. The function words are also high-frequency and help to maintain the cohesion of the sentences and the verbs and pronouns help to uphold the semantics. It is also in line with the perception of cognitive architecture introduced by Likourezos (2021), who argues that only a few new pieces of information can be processed by working memory simultaneously. Digital communicators can decrease the complexity of the syntactic components of a message and avoid cognitive overload, thus allowing important information to be easily digested by the sender and receiver. Simply put, the simplification in the data is more a practical usage of cognitive load management in real-life language use taking the theoretical implications of classroom and pedagogical settings to the sphere of online communication.

Moreover, the syntactic patterns that are selective in the Gen Z messaging can be explained within the framework of the difference between primary and secondary knowledge (biological) (Sweller, 2011). In a familiar social digital communication, communication is considered primary knowledge processing and this can be done with little conscious effort. Nevertheless, organization and clarity of messages, in particular, the delivery of new information, need to be addressed to secondary domain-specific processing, which adds extra cognitive burden. The balance between these two forms of processing is probably the reason why users use a combination of short sentences in their usual interactions and longer sentences with more details in the delivery of new or complex information. That is, the organizational trends that are observable in the data are adaptive strategies to maximize allocation of cognitive resources in accordance to the tenets of Cognitive Load Theory.

It can also be discussed due to the high presence of punctuation marks (PUNCT = 935). Punctuation is a cognitive scaffold device that allows the reader to effectively decode the sentences and decreases the unnecessary cognitive load related to syntactic relation interpretation (de Jong, 2010). Likewise, the minimal number of numerals (NUM = 77) used and interjections (INTJ = 35)

show that the content that may add unjustified load is kept to the lowest possible except in circumstances where it is needed. Mathematically speaking, these results resemble the suggestions found in the educational literature in that less extraneous load leads to improved understanding and simplified process of transferring information held in working memory to long-term memory (Clark and Kimmons, 2023; Sweller, 2011).

In general, the NLP analysis shows that the concept of syntactic simplification and selective engagement with lexicon could be observed in Gen Z online communication. These tendencies may be explained by the practical applications of the principles of Cognitive Load Theory when the management of intrinsic and extraneous cognitive load is controlled by the user, and the optimization of germane cognitive processes takes place. The use of short sentences, excessive use of pronouns and verbs, low use of complicated syntactic constructions, and purposeful punctuation are all indicative of the strategies to achieve maximum communicative efficiency and the minimum amount of cognitive effort. This interpretation can be compared to previous studies that revealed the inherent limitation of working memory that determines how people use language in communication situations that occur in real-time (Sweller, 1988; Likourezos, 2021). In addition, this result corroborates the larger argument that simplification in language application is not a weakness but rather an adaptation mechanism based on cognitive architecture and memory limitations.

CONCLUSION

The purpose of this research was to consider the trends of syntactic reduction in Generation Z online communication with the help of computational linguistic analysis and Natural Language Processing (NLP) tools. The study involved analysis of a dataset of 160 Reddit posts (a discussion thread in the r/Law School community) (Additional-Reserve66, 2025). These remarks are informal web communications that users have created most probably, with the Generation Z group. In order to find the patterns in the structure that define the modern digital communication, the study examined various linguistic characteristics, such as sentence length, part-of speech distribution, and frequencies of words using Python and NLP libraries, including NLTK, spacey, Stanza, and Text Blob.

According to the analysis findings, Gen Z digital communication exhibits observable tendencies of syntactic simplification. The dataset had 10,476 words in total and the mean number of words in a comment was 65.47. Analysis of the sentence's length showed that there was a significant range in the structure of the sentences, as the sentences had a minimum of two words and a maximum of 119 words with the means being 23.28 words. However, longer sentences might be used, though the facts indicate that the users often utilize shorter and less complicated constructions. The part-of-speech distribution also follows this observation as verbs, pronouns and auxiliary verbs were more likely to occur than a descriptive word like adjective or proper nouns. On the same note, the highest frequency words in the data were common functional words like the, to, a, I, and, and which means that communication is based on simple grammatical patterns that support the transfer of clarity and efficiency.

The rationale behind the study is the growing importance of digital communication platforms on the daily language practices, especially of the younger generations. The Internet can generate vast volumes of naturally occurring linguistic data in online forums like Reddit where users communicate informally using written communication methods. This aspect of language functioning in such settings may be of use to investigate the impact of digital interaction on modern syntactic frameworks. Using the methods of computational linguistics on such data, the given study attempts to offer empirical evidence on the issue of communication patterns development in a digital setting.

This research also substantially finding is expected based on Cognitive Load Theory (Sweller, 1988) that indicates that there is a limit to human cognitive processes, which is the capacity of working memory. The effect of this is that in high processing speed digital communication, users can simplify the structure of sentences naturally to reduce the cognitive load and make the transfer of meaning within sentences more efficient. The usage of frequent function words, prevalence of verbs and pronouns, and the general scarcity of complex syntactic constructions help to justify the assumption that communicators are more likely to focus on easier to realize and interpret structures on the fly.

The study would be relevant to the area of computational linguistics and computer assisted discourse analysis by showing how NLP tools may be applied to the analysis of syntactic patterns

in online communication in its natural form. Combining computational analysis with the Cognitive Load Theory, the research demonstrates that the linguistic behavior in the digital space may be explained by the principles of cognition that govern how people cope with the processing load in communicative interactions.

Moreover, the results indicate that syntactic reduction in Gen Z online communication cannot be regarded as the loss of linguistic skills and can be seen as an adaptive mechanism influenced by the communicative and cognitive requirements of online communication. All in all, the present paper yields an understanding of how the modern digital discourse is an indicator of latent cognitive limitations and also shows the usefulness of the computational methods in the current linguistic studies.

REFERENCES

- Additional-Reserve66. (2025). *I love watching my “test anxiety” friends leave an exam 2 hours after I was required to finish mine* [Online forum post]. Reddit. <https://www.reddit.com/r/LawSchool/comments>
- Awati, R., Gillis, A. S., & Lutkevich, B. (2025, October 28). What is natural language processing (NLP)? *TechTarget*. <https://share.google/2jXZOdeLqt5MKMu6W>
- Azaz, M. Z., Shahid, A., Rashel, M. M., & Nelson, J. C. (2024). The digital linguistic edge: How it is transferring the study of language. *Educational Administration: Theory and Practice*, 30(5), 14544–14558. <https://doi.org/10.53555/kuey.v30i5.6891>
- Beck, M. (2025, April 18). Gen Z is destroying literacy with netspeak and leetspeak. *The Colgate Maroon-News*. <https://share.google/dvS82A47eiIHsbcVY>
- Box News. (2025, February 5). What is natural language processing (NLP)? *Box News*. <https://share.google/cg9zKPMrMX7hDZzGd>
- Chandrasekar, R., Doran, C., & Srinivas, B. (1996). Motivations and methods for syntactic simplification. *Research on Language and Computation*, 4(1). <https://doi.org/10.1007/s11168-006-9011-1>
- Clark, C., & Kimmons, R. (2023). Cognitive Load Theory. In *EdTechnica: The Open Encyclopedia of Educational Technology* (pp. 109–115). EdTech Books. <https://doi.org/10.59668/371.12980>
- Cole, S., & Holdsworth, J. (n.d.). What is NLP (natural language processing)? *IBM Think*. <https://share.google/20YXKsyf8dMHu6JZ4>
- Crystal, D. (2009). *Language and the Internet* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511487002>
-

- Dawn News. (2025). The vivid language that Gen Z speaks. <https://share.google/e8TladihjlyvHnhum>
- de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, 38(2), 105–134. <https://doi.org/10.1007/s11251-009-9110-0>
- GeeksforGeeks, Sanchhaya Education Private Limited. (2025, July 23). History and evolution of NLP. *GeeksforGeeks*. <https://share.google/bVxIUywOCqwFDkOq8>
- Ivchenko, O. (2022). Impact of text simplification on understanding. In *Challenges of trustable AI and added-value on health* (p. 634). <https://doi.org/10.3233/SHTI220546>
- Jabeen, S. S., & Nawaz, H. I. (2025). Morphological innovations in Gen Z language and their impact on English language usage. *Journal of Applied Linguistics and TESOL (JALT)*, 8(1). <https://share.google/mNmPnVXRfSkcTaHcG>
- Javeed, S. (2025, June 11). The use of Generation Z vocabulary and the English language proficiency among Grade 11 students at Conner Central NHS. *American Journal of Educational and Social Science*, 51(6), 1170–1187. <https://doi.org/10.9734/ajess/2025/v51i62066>.
- Likourezos, V. (2021, March 3). *An introduction to cognitive load theory*. <https://share.google/xtI8EfuS8gEnsMUA6>
- Lorenzo, J. P. (2025). The use of Generation Z vocabulary and the English language proficiency among Grade 11 students at Conner Central NHS. *Asian Journal of Education and Social Studies*, 51(6), 1170–1187. <https://doi.org/10.9734/ajess/2025/v51i62066>
- Marzona, Y. (2025, June). Language change in the communication of Generation Z. *Jurnal Ilmiah Langue and Parole*, 8(2), 153–161. <https://doi.org/10.36057/jilp.v8i2.734>
- Panajaitan, L. L., & Patria, A. N. (2024). Social media and language evolution: The impact of digital communication on language change. *International Journal of Linguistics, Literature & Translation*, 7(12), 53–57. <https://doi.org/10.32996/ijllt.2024.7.12.8>

-
- Polzin, A. (2022, March 23). Natural language processing – A brief history. *HIVE: Big Picture Blog*. <https://share.google/Yc7UuoKaDTuqEqh0y>
- Siddharthan, A. (2006). Syntactic simplification and text cohesion. *Research on Language and Computation*, 4(1). <https://doi.org/10.1007/s11168-006-9011-1>
- Sitohang, D. P., & Ristia, B. A. (2025). A semantic analysis of slang words used by Gen Z on social media. <https://share.google/B0EuSJDQDIiWa3TO0>
- Stăncescu, N. A. (2025). The evolution of language in the digital age. *Social Economic Debates*, 14(1). <https://share.google/oiTCReHbzCZiWnoSV>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 37–76). Elsevier Academic Press. <https://doi.org/10.1016/B978-0-12-387691-1.00002-8>
- Telaumbanua, Y. A., Crsitin Zendrato, N. E., Nazara, B. T., & Harefa, E. C. (2024, August). Analyzing the language style of Generation Z teenagers in their Instagram posts. *JELITA*, 5(2), 588–600. <https://doi.org/10.56185/jelita.v5i2.790>
- Tufail, S., Asgher, M., & Ali, I. (2024). Deciphering the digital vernacular: A study of Generation Z slang on social media platforms. *Pakistan Languages and Humanities Review*. [https://doi.org/10.47205/plhr.2024\(8-II\)75](https://doi.org/10.47205/plhr.2024(8-II)75)
- Yeasmin, S. (2025). Language as a tool for defining generations: A study on Gen-Z lingo. *International Journal of Advanced Research*. <https://dx.doi.org/10.21474/IJAR01/20328>
-