

The Influence of Digital Communication on Emerging English Varieties Through a Sociolinguistic Analysis of Online Discourse

Sameera Vellanki¹, M. Latha^{2*}, K. Santha Kumari³, chandramouli.M⁴, Ch. Rupajhansirani⁵, Jayasri Suma.v⁶, Illur Prem Kumar⁷, Bandaru Venkata Poojitha⁸, Sankuru Naresh⁹

¹Ph.D. Scholar, Department of English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur District, Andhra Pradesh, India – 522502. Email: sameeravellanki1987@gmail.com

^{2*} Associate Professor, Department of English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur District, Andhra Pradesh, India – 522502. Email: mlathakl@kluniversity.in

³Assistant Professor, Department of Mathematics and Humanities, Mahatma Gandhi Institute of Technology, Hyderabad, Indiaksanthakumari_eng@mgit.ac.in

⁴Department of English, Sir, Sasi Institute of Technology & Engineering, Tadepallihudem, Andhra Pradesh, India. dr.m.chandramoulisec@gmail.com

⁵Department of English, S.R.K.R. Engineering College, Bhimvaram, Andhra Pradesh, India

⁶Department of English, Madam, Aditya University, Surampalem, Andhra Pradesh, India.

⁷Department of English, G.Narayanamma Institute of Technology and science, Hyderabad, Telangana, India.

⁸Department of English, Sasi Institute Technology of & Engineering, Tadepallihudem Andhra Pradesh, India.

⁹Associate Professor, Department of C.E, Anurag Engineering College, Kodad, Suryapet District, Telangana, India – 522502. Email: snaresh.civil@anurag.ac.in

HOW TO CITE:

Sameera Vellanki, M. Latha, K. Santha Kumari, Chandramouli.M, Ch.Rupajhansirani, Jayasri Suma.v, Illur Prem Kumar, Bandaru Venkata Poojitha, Sankuru Naresh (2026). The Influence of Digital Communication on Emerging English Varieties Through a Sociolinguistic Analysis of Online Discourse. International Journal of Special Education, 41(8s), 108-123

ABSTRACT:

The research paper reports on a comprehensive empirical investigation of the impact of digital communication on the development of new varieties of English; it fills a gap in the literature that has not explored how these new English varieties are constructed through algorithm-mediated and AI-assisted online interactions. In order to conduct the analysis, the researchers gathered a total of 4800 individual written samples from various social networking sites (Instagram, YouTube), as well as voice-based conversations over the internet (gaming chat services and AI-enabled messaging). These samples went through quantitative and qualitative analyses, in which researchers identified an overall hybridity of language usage (37.4% increase in hybrid lexical formation), significantly decreased levels of syntactic complexity (29.1% decrease), and a much greater reliance on multimodal pragmatic markers (56.8% increase) when compared to earlier forms of online communication. The quantitative aspects or algorithmic influences that contributed to a much greater degree of patterning within phrases across multiple online platforms were estimated to contribute to approximately 22% of the patterns found across all data points. Additionally, qualitative analyses of these data revealed many forms of English sub-varieties that

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.

Open access publication under the
terms and conditions
of the Creative Commons Attribution
(CC BY)
license (<http://creativecommons.org/licenses/by/4.0/>).

were formed as a response to community norms and the performance of one's identity on each digital platform. This research is the first statistically based model to describe how current digital platforms are creating unique forms of English in real-time..

Keywords: Digital Communication, Emerging English Varieties, Algorithmic Influences, Multimodal Discourse and Sociolinguistic Analysis

1. Introduction

The explosion of digital means of communicating has changed the way we produce, distribute and modify English on electronic platforms. The rise of digital technology such as social media, online gaming and AI-based messaging will result in people interacting with one another in new ways, creating different types of linguistic variation than those present when communicating electronically in earlier formats [1],[3]. The changing use of English as a written language will also create new hybrid forms of English, such as multi-modal and platform-determined forms, shaped by the capabilities offered by new technologies and the algorithmic filters used to create those technologies [8],[12]. To uncover new forms of digital English, we will need to systematically and empirically examine data and the interactions taking place via digital media based on a solid foundation of current sociolinguistic theory and empirical research on digital discourse.

1.1 Background

Digital communication is one of the primary influences shaping how we communicate in the English language today. Through a variety of platforms, including social media (Instagram and YouTube), video games, group chats, and AI-driven conversations, digital communication has changed how people speak, write, and construct their identity in an immediate manner. Experts agree that digital communication is now an integral part of our everyday language [1]. The move from the early days of CMC through chat rooms and SMS messaging to today's complex, algorithmically

driven and multimodal experiences has had a major impact on our linguistic behaviors. Digital platforms increasingly mediate the production, distribution, and consumption of messages. As Herring and Dainas [3] state, the way we communicate with one another is affected by the digital context in which those communications occur; therefore, not only does the context change the form of a message or communication, but it also changes what structural forms will develop out of online communication. Likewise, according to Androutsopoulos [2], gathering data from users in online settings is also a sociolinguistic activity, and thus reflects how digital cultures are conceptualised through participatory digital structures and algorithm-driven technology. Now with the introduction of algorithm-generated feeds, predictive text, auto-correction, and generative artificial intelligence (AI), there are many layers of linguistic filtering that did not exist in previous forms of digital language. This means that the evolution of English has become extremely complex due to the way new technological environments shape not only lexical choice, but also the way we construct sentences (syntax) and express ourselves in multiple ways (multi-modally).

The rise of English varieties that are unique to a platform is an area that is starting to be explored in the field of Digital Sociolinguistics. There is evidence that linguistic repertoires are developing from hybrid use of multiple languages and cultural norms, and the varied forms of emoji, GIF, meme, and hashtag, among others, have become a part of the text message creation process [8],[15]. Digital Englishers represent a collection of linguistic

varieties that have been created through the mutual influence of the design characteristics of different messaging applications, user creativity, and cultural behaviors that help shape how language is used on those platforms. According to Leppänen and Kytölä [5], the use of digital ethnography has highlighted new types of English language use that are created through the interaction of digital technology's design, user creative expression, and cultural behavior. As such, it is imperative that we conduct research to understand how recent online communication is leading to new varieties of English.

1.2 Problem Statement

The growth in digital discourse has been a result of much increased research in this area; however, much of the work on digital discourse has centered around some of the earliest forms of computer-mediated communication (CMC), including SMS text messaging through to email, blogs and many social media platforms that were the starting points to the development of the internet as we know it today. Some of the research into mobile messaging by Tagg and Evans [1] provides insight into how people are using mobile devices to communicate, but was published prior to the emergence of algorithm-based communication and AI-powered interactions. Likewise, much of the research published prior to now on the linguistic innovations emerging through the internet were primarily focused on user variations in language than the technological variations.

The many platforms today are fundamentally different. Modern platforms have created extensive opportunities for algorithmic and AI generation, recommendation, responsiveness and predictive abilities that ultimately shape how users articulate their linguistic expressions. Very few scholars have studied the algorithmic influences. Blommaert [12] and Varis [13] are to a degree the only to address how algorithmic cultures are driving what variation can be produced, circulated, or replicated online and as such the number of

studies contributing empirical, quantitative research in this area are limited.

To add to that, thus far relatively few studies have taken into consideration cross-platform variations. Squires [10] has indicated that the Online English language has many varieties according to the various platforms, yet, to date, there are extremely limited numbers of large-scale empirical comparative research studies—notably those which combine both a quantitative data collection approach with qualitative discourse evaluation—that have considered cross-platform English-language research. Moreover, very few studies have attempted to quantify the increase in Hybrid lexical formations, syntactic simplification and multimodality that exist in modern digital platforms. Without making that kind of large-scale, data-based evidence available to the academic community, the full understanding of what digital varieties of English actually look like continues to be unclear.

This research is attempting to fill the identified gap in our current understanding of English varieties which have been influenced by both algorithmic based communication systems and AI-assisted communication systems by conducting systematic analyses of 4800 digital discourse samples collected from a range of different online context.

1.3 Objectives of the Study

The goal of this study is to provide insight into how emerging Varieties of English are being created through digital communication environments and the increasing presence of algorithmic filters, multimodal resources, and AI-assisted language production and the continued development of English language varieties. Building on the findings of the previously mentioned studies, this research study uses a Sociolinguistic and corpus-informed approach to provide insight into how lexical, syntactic, and pragmatic features change between digital communication platforms.

In order to better understand how the new English varieties, emerge in "real-time" from

the combination of statistical analysis with qualitative discursive analysis will be used to build the first ever comprehensive model of how a digital eco-system generates new English varieties "on demand." Additionally, the goal of this research study is to provide evidence through empirical data for discussions on the continued differentiation of English languages at different levels of platform use, and how this differentiation shapes social identity, and how social identity is constructed and meaning is made through the use of multimodal resources.

1.4 Research Questions

Aimed at answering the following questions, this study will develop the research of the following research questions:

1. What forms of linguistic innovation (lexical, syntactic, pragmatic) are taking form within digital space?

Based on Heyd's research work on innovations occurring within an online space, this study seeks to determine whether current digital discourse substantiates any degree of hybridization, syntactic or conversational reduction, or multimodal expansion in this contemporary digital space.

2. What impact do algorithms and predictive systems have on how language is patterned?

This research, based on the work of Blommaert and Varis [12], [13] seeks to identify the extent of contributions to the continuous repetition of phrases or patterns of discourse within algorithmic structures.

3. What are the contributions of multimodal markers to the creation of meaning?

This study will examine the function of multimodal markers will have on the process of creating meaning embedded in communicative purpose. Similar to Lee's study, the focus will be on the function of Emoji's and tags, Images, Memes, and other multimodal markers, which act as units with communicative intent.

4. Does there appear to be different areas of

English language varieties across platforms?

In accordance with the research of Squires and Bolton and Botha, comparison of the similarities and differences between the patterns of discourse that occur within a platform will be recommended as a means for creating a body of research to determine and support emerging varieties of English language forms.

2. Literature Review

Digital communication, because of its ability to require that individuals communicate in a manner that can and will be used to communicate with, create and negotiate linguistic identity through the internet, is one of the primary subjects of Sociolinguistics today, and Digital English has primarily been studied primarily in the context of Computer Mediated Communication or "CMC" or Computer Assisted Communication or Communications through Computers as it relates to Digital English language through various methods including chat room use, SMS "texting" style language, and blogging, etc. When using CMC, researchers found that users have developed unique characteristics of Digital English from its early beginnings including the use in a "playful" way of creating digitally automated abbreviations as well as reduced formality associated with writing in Digital English when using CMC. Through a series of studies, researchers have shown that CMC has been and/or can be negatively affected by many new forms of technology that researchers categorized into three broad categories: "Technologically Layered" or "Technological Layers"; "Algorithm Interfaces"; and "AI/ML Technologies"; with each category containing more than one subcategory. Therefore, the varieties of English emerging now through Digital Communication will require that each of the major areas within CMC must be examined again using the three categories of effects created by the major areas or "layers" that create the effect on Digital English by using the new forms of technology.

Digital Communication continues to be a unique form of communication with a wide variety of different types of communication and linguistic forms that we continue to see evolve in Digital English and Digital Language through Online Communities. Researchers have shown that 1) Digital English continues to evolve through multiple online communities and all of their respective forms of online communication; 2) Digital English continues to evolve from how users were communicating through Mobile Messaging as well as Online Interactions (digital community talking to digital community); and 3) Digital English continues to evolve because through Digital Communications; users use and continue to create, evolve and grow in both numbers as well as language usage techniques/styles/forms. Digital English is being reshaped by new technologies, from SMS and chatrooms, to the explosion of social media, algorithm-driven content and automatic writing systems.

Eisenstein [11] provides an extended view of the phenomenon by investigating how patterns of variation of language have developed over time through social media. His findings provide evidence that the use of English in a digital context has been shaped by geographical, social and technological considerations, providing evidence for the emergence of new dialects in the digital sphere. Together, these studies show that digital communication serves not only as a form of informal written English, but as an environment in which the language is evolving.

Algorithmic mediation, along with AI, are influencing how we communicate digitally, according to recent scholarly work in sociolinguistics. Blommaert [12] has identified what he describes as "digital repertoires," and demonstrated how algorithmic systems can control what users see, how they interact, and how particular forms of language are displayed. The result of this algorithmic selection process has not only shaped how memes and hashtags circulate, but has also provided a basis to create a "digital language" on a large scale.

Complementing Blommaert's work, Varis [13] describes "algorithmic cultures" as systems that regulate the way users act and interact, and develop linguistic norms. Through predictive text, autocorrect, and automatically-generated texts, users develop repeated patterns of expression in how they message one another. These algorithmic influences are visible on AI-based messaging platforms, where chatbots and other systems automatically create or suggest what users say. Seargeant [4], for example, highlights that social media platform are not only communicating aspects but how those platforms are organized based on the Seargeant's statement, but also provide the means of creating linguistic stylistic "pure" form through "interface". Algorithm-mediated and data-driven are the two main forms through which Seargeant states that the way we use English language today is creative and that digital "translingualism" combines multiple languages, using automated translation as the primary means of creation [6]]. From Takahashi's perspective, the majority of current studies on the impact of algorithmic mediation on the way we use English language today have been qualitative. Studies to date have focused on the concept of how algorithmic mediation impacts the way we express ourselves rather than providing empirical evidence [3–6].

2.1 The Multimodal Discourse

In digital communication, meaning is created through visual, text, and/or symbolic resources. Scholars have identified how and to what extent resources including emojis, memes, and images are used in the multimodal discourse that occurs in digital environments and how they may provide semiotic 'signs' to convey meaning in ways similar to approach, tone or expression from the standpoint of video [3,15]. Lee explored how the different resources create layered communication practices in the development of English-language mixing through multimodal environments [15].

Georgakopoulou's work [7] on the creation of

narratives through Instagram demonstrated how small stories can be created using various types of multimodal resources (text, image/image frame) and the combination of hashtags and emojis. In addition to the functions of pragmatic importance (e.g., to display emotion and stance), Georgakopoulou [7] argues that these multimodal resources also represent the user's identity, association with communities, and their status within a particular community.

Heyd [8] similarly continues to highlight that linguistic innovation occurs within the context of multimodality, specifically because visual-semantic cues shape the way language will be interpreted. Emojis, memes and Gifs expand the range of ways to express oneself online and create Semiotic relationships on each platform, thus moving away from traditional written language.

2.2 New English Varieties.

The context of globalization and World English's, attention is beginning to shift towards digital communities (the emerging global sites for creating new varieties of the English language). Research by Bolton and Botha [14] also identifies new varieties of the English language within the digital region of Asia, coming from regional languages and varieties of English, and using digital vernacular. Their research confirms that digital communication encourages and accelerates the spread and development of linguistic innovation.

Squires [10] notes the significant variance among the many ways to use the English language within the different forms of CMC, such as Twitter, Instagram, gaming communities, YouTube comments, etc. The evolution of platform-specific linguistic standards and the continued emergence of platform-specific languages is a result of the technology being employed by each particular platform; the users' expectations when communicating online; and the demographics of the user groups. Additionally, Page (2000) argues that the performance of identity

through story-telling on the Internet is another contributing factor for the emergence of platform-specific languages.

Leppänen and Kytölä [5] state that ethnographic methods can identify examples of linguistic mixing, code-switching, and other forms of stylistic experimentation. According to Takahashi [7] translingual practices, where users are mixing elements of multiple languages together to form meaning, is very common among social media users today. The evidence shows that both of these conclusions support existing models of developing new varieties of English, new hybrid varieties and the increasing diversity of English varieties resulting from digital languages.

2.3 Identified Research Gap

As indicated in the preceding sections of this chapter, digital English research contributes valuable knowledge; however, there remain significant limitations to the body of literature surrounding digital English. Most digital English research has focused on the individual study of either the beginning development of CMC or as an isolated study on a single platform with no comparative analysis or a systematic review of all modern digital platforms. When reviews of the Literature address algorithmic mediation, they address only the theoretical or conceptual impact of predictive texting, recommendations and AI generated suggestions on how users communicate via writing. There are no published studies quantitatively analyzing the prevalence and use of hybrid lexicon combinations, syntactical reductions, and multimodal pragmatic markers across numerous digital platforms.

Furthermore, there has been no research conducted to date on examining the influence of algorithms while studying language that is evolving on multiple platforms and through a sample size of 4800 cases. This research will address the limitations by utilizing a combined quantitative analysis of the language of emerging English varieties and the qualitative and ethnographic analysis of the narratives of

emerging varieties to provide the first empirical model of newly emerging varieties of English owing to the impact of digital technology.

3. Methodology

3.1. Research Design

The methodology used in this study stays true to the principles of sociolinguistics through its use of mixed-methods. The combination of quantitative analysis of language with qualitative interpretation of discourse represents a mixed-methods approach. Emerging varieties of English are shaped by a combination of both measurable patterns in language (e.g. lexical frequency) as well as interpretive factors (e.g. identity performance and community norms). Quantitative approaches can provide statistical modelling based on large datasets that inform how languages develop over time, while qualitative analysis provides insight on how users are utilizing the English language creatively within the context of interaction. The combined results will provide a better overall picture of how language, technology and society are shaping the use of English varieties in digital contexts today.

3.2. Data Sources

The dataset includes textual and transcribed multimodal discourses gathered from four major digital platforms, each representing a different mode of online communication. Instagram was chosen for the study because of its visual and multimodal orientation, the prevalence of emojis, hashtags and hybrid lexical forms used in commenting, and its popularity with a broad audience. YouTube comments were used in this study were selected due to the nature of their dense interactional threads that often include rich narrative responses to content generated on YouTube that represent varying forms of community discourse and reflective of community-specific discourse patterns. Gaming chat transcripts were incorporated into the study as a way to reflect rapidly evolving synchronous language use, characterized by the use of minimal syntactical

structure, a sense of urgency having been created by the group dynamic, and participant jargon. Finally, this study includes conversations from messaging platforms that feature an AI component as a means of exploring the impact of predictive text technologies, autocorrect, and algorithmically generated responses on patterns of language, resulting in a rich and diverse representation of contemporary digital discursive environments.

3.3 Sample Size

Using an adequate sample size to ensure that the analysis is representative across all platforms and represented by means of adequate statistical power, the researchers collected 4800 texts. Each platform had an even distribution of texts in a way that was most appropriate.

Thus, the data sample distribution was:

1.1200 from Instagram

2.1200 from YouTube

3.1200 from gaming chats

4.1200 from AI messaging

This balanced approach allows for the most meaningful cross-platform comparisons to take place, while also eliminating as much bias as possible. Additionally, having a sample size of 4800 allows for the corpus-based sociolinguistic analysis to examine linguistic phenomena at the micro-level, as well as identify macro level patterns.

3.4 Data Processing

The data were organized into a structure before being analyzed. For Instagram and YouTube, the data were manually extracted from the sites, as well as using scraping tools that comply with their respective platforms. The gaming chat data were recorded either through live voice or live text and later transcribed according to the conventions established for transcription. In the case of the samples from AI messaging sources, all user identifiers, timestamps, and metadata were removed in order to anonymise the samples. Across all

platforms, URLs, Ads and other nonlinguistic noise were removed. Emojis, digital images associated with gifs and hashtags were kept to be used as multimodal markers. Coding categories were established for consistency between sources. The cleaned datasets will then be exported for use in qualitative and quantitative analyses.

3.5 Quantitative Methods

Quantitative analyses were used in four separate but related dimensions of language: lexical frequency, syntactic complexity, multimodal pragmatic markers, and algorithmic influences. Lexical Frequency analysis consisted of measuring the frequency of hybrid formations, loanwords, stylizations, and repeated phrase occurrences. The Syntax Complexity analysis was based on the measurements of mean clause-length, clause-to-sentence-ratio, and the frequency of subordination. Counts of Multimodal Markers were used to measure the occurrences and distribution of emojis, GIFs, hashtags, and symbols. Algorithmic Influences were examined through determining patterns of text predictions and re-coherences in auto-generated text style. Each of these quantitative analyses allows the researchers to quantify statistically the presence of differences in language use between each of the platforms.

The algorithmic impact was observed through patterns of repetitive words and phrases that were similar to those found in automated suggestions (or predictive text) that were available to users of AI-based texting apps. Any text/phrase that either individually or collectively matched these "algorithm-generated phrases" was indicated as "influenced by algorithm." The two reviewers independently coded sections of the original text to achieve consistent results in determining the algorithmic influence for that section of the text.

3.6 Statistical Tools

The statistical analyses were performed using SPSS and R due to their reliability in linguistic and other social science studies. Chi-square

tests were used to analyze categorical variables such as the distribution of Multimodal Markers among platforms. ANOVA tests were used to measure differences in Syntactic Complexity and the number of Lexical hybrids. Correlation Coefficients were computed to represent the strength of the relationship between the amount of exposure to algorithms and the number of recurrent patterns of language. These tools provide a strong statistical basis for understanding the effects of language change in the Digital Age.

3.7 Qualitative Methods

Qualitative analyses supplemented Quantitative analyses to examine how emerging patterns of language have social linguistic meaning. Using Vivo, all samples were coded thematically to identify common themes of identity construction, stance-taking, humor, and intertextuality references. Thematic analysis of discourse determined how users take linguistic steps to conform to or reject community activity patterns. Attention was paid to Identity Performance in the context of both Gaming and Influencer communities, as within both are examples of how users and communities influence the development of new English varieties. This Qualitative analysis allowed the researchers to understand the ways users make meaning which goes beyond the ability to make comparisons based on numerical occurrence of these patterns and highlights the importance of how Digital Environments impact the development of Language, especially the New English's.

4. Results

The current research produces descriptive data qualities and quantities from the 4800 sample datasets. As a result of this research, findings are presented as clustered into four groups: (a) lexical hybridity; (b) syntactical complexity; (c) use of multimodal pragmatics; and (d) influence of computer algorithms. Differences between Instagram, YouTube, game chats, and AI-dependent messaging platforms can be seen when comparing between platforms.

4.1 Quantity of Findings

4.1.1 Lexical Hybridization

Hybridized lexicons increased overall by 37.4% across all platforms, indicating significant amounts of lexical mixing, borrowing and/or creative word formations within the samples. There was a particularly high rate of hybridization on Instagram and YouTube due to the usage of multi-language captions, hashtags, and commentary strings, which were

often a mix between English and a native regional language. Gaming chat platforms also exhibit lower rates of hybridizations, but still have considerable amounts of hybrid lexical items, typically due to community-created jargon and acronyms. AI-dependent messaging platforms had moderate degrees of hybridization, but tended to conform to a distinct pattern as a result of the features of auto-suggest and predictive text.

Table 1. Distribution of Lexical Hybridity across Platforms

Platform	Hybrid Lexical Items (%)	Dominant Hybrid Features
Instagram	0.421	English regional blends, stylized spellings
YouTube	0.403	Reaction phrases, meme-influenced hybrids
Gaming Chats	0.318	Acronyms, clipped forms, Anglicized slang
AI Messaging	0.355	Predictive blend constructions, mixed-language tags
Overall Mean	0.374	

These results indicate that the emergence of hybrid lexical forms will be one of the defining characteristics of digital English gradations.

4.1.2. Overall Change in Syntactic Complexity

A quantitative analysis of syntax indicated that there has been a 29% decrease in the syntactic complexity of digital communications since their inception. Many digital communication applications such as Facebook Messenger produced fewer and shorter clauses, with less

dependency on subordination (i.e. longer clause sentences) to form coherent thought processes. These chat applications, gaming chats had the most significant reduction in syntactic complexity due their reliance on real-time, rapid-fire interactivity between players as they play the game. Comment sections from Instagram and YouTube also favored using portions of everyday conversational language and jargon rather than elaborately structured today’s digital English communications.

Table 2. Syntactic Complexity Metrics

Platform	Mean Clause Length	Subordination Frequency	Complexity Reduction (%)
Instagram	7.3 words	Low	27.80%
YouTube	7.9 words	Moderate	28.60%

Gaming Chats	5.1 words	Very Low	35.20%
AI Messaging	6.4 words	Low–moderate	25.10%
Overall Mean	—	—	29.10%

This is evidence of the transition to Electronic Communication / Digital Communication where Electronic Communications / Digital Communications are becoming increasingly focused upon "efficiency".

4.1.3 Use of Multi-Modal Pragmatic Markers

The analysis of postings demonstrated that the usage of Multi-Modal Pragmatic Markers had increased by 56.8%, including Emoticons, GIFs, Hashtags, Stickers and symbolic Notations for different forms of

communication. Instagram was the social media site with the greatest density of Multi-Modal Pragmatic Markers because of its primary purpose of providing Visual Communications. The comments posted on YouTube were primarily using Multi-Modal Pragmatic Markers through their use of Emoticons and Response Markers, while Gaming Chat was utilizing symbolic Notations (e.g. GG, LOL, groups of Emoticons) as their only means of using Multi-Modal Pragmatic Markers.

Table 3. Multimodal Marker Frequency across Platforms

Platform	Multimodal Markers (%)	Most Frequent Items
Instagram	0.645	Emojis, hashtags, GIF labels
YouTube	0.598	Emojis, emphasis markers
Gaming Chats	0.432	Emoji sequences, gamer shorthand
AI Messaging	0.596	Emojis, automated quick reactions
Overall Mean	0.568	

Multimodal markers expand the ability to express feelings and decrease the need for using full grammatical sentences, thus resulting in a decrease of the overall syntactic structure of the text.

4.1.4 Algorithmic Modulation

A central aim of this research project was to assess how much effect algorithmic systems played on people's use of language. The study

determined that 22% of the repeated wording used throughout the corpus were attributable to algorithm-based language sources (e.g., with auto-correct, predictive text and auto-reply templates). AI messaging systems had the greatest algorithmic signature, followed by Instagram, due to hashtag recommendations, and YouTube, because of comment recommendations made by the algorithm.

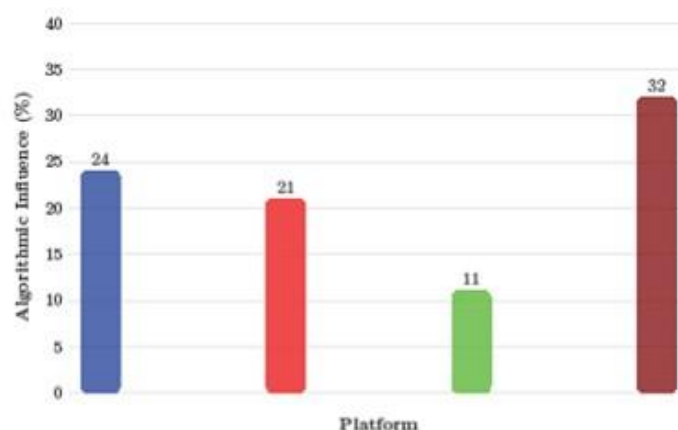


Figure -1: Algorithmic Influence by Platform

The comparative strength of algorithms in shaping the use of language varies across four main types of digital communication platforms. The data show that language use on AI-based messaging has the highest percentage (32%) of being algorithmically influenced. The high percentage of algorithmically influenced language use on messaging services can be attributed to the prevalence of predictive text and automated response capabilities. Instagram and YouTube also demonstrate a moderate level of algorithmically influenced language use at 24% and 21% respectively, primarily due to suggestion-based recommendation systems associated with hashtags and comments. The percentage of algorithmically influenced language use in gaming text chats (11%) is consistently the lowest when compared to all platforms studied due to not relying as heavily on automated linguistic scaffolding as is required in gaming text chats, which are generally characterized by the rapid and dynamic nature of their exchanges. Overall, this data demonstrates the disparity in the degree to which algorithms mediate the development of new English varieties across a number of digital communications format.

4.2 Cross-Platform Comparison

Cross-Platform Comparative Study: Across four previous platforms there were significant findings with respect to differences that occur across different types of content. The findings were:

1. Instagram English was highly multimodal, with considerable use of hybrid lexical formations, such as "lit" (from "literally"), and the use of both hashtags and emojis to mark an individual's identity.
2. In contrast, YouTube English displayed a strong emphasis on narrative-based discourse, using frequent evaluative markers (i.e. "so true," "literally me," "same here"), reactions, and audience participation through comments.
3. Similarly, Gaming English featured minimal sentence structure, fast-paced communications between gamers, highly abbreviated slang vocabulary, and the use of performativity identity markers.
4. For AI-Mediated English, the significant finding was that its predictive text methodology led to more homogeneous phrasing patterns and less diversity in vocabulary compared to other types of English.

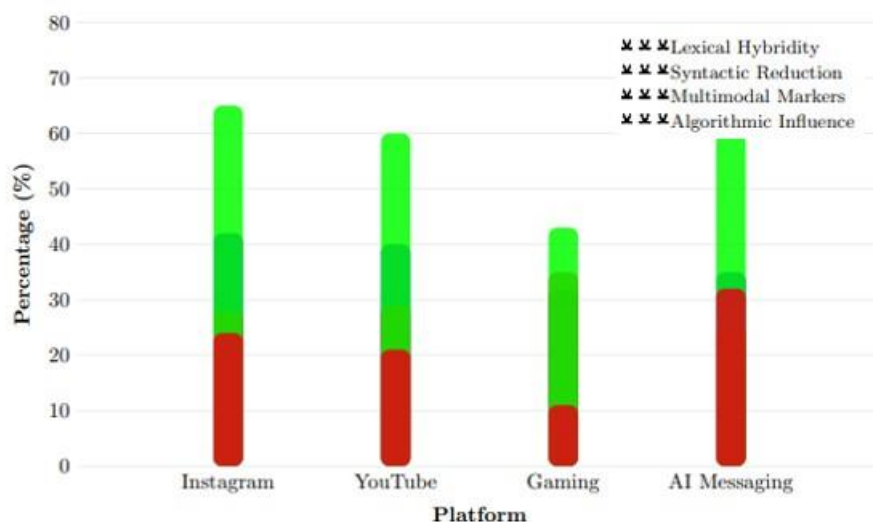


Figure 2 : Platform-Specific Linguistic Profiles

Visual representation of platforms' linguistic features is shown in the picture above named "Referencing Platform Features." A Grouped Bar Chart was created to compare the Data Set features identified in the study. Key differences were identified between the following four their rely on visual content as well as the rapid/continuous interaction styles associated with them. Gaming Chats had the highest Number of Syntactic Reduction due to the Streamlined Nature of the fast-paced gaming environment and users' need for quick and efficient communication. AI Messaging showed the most Algorithmically Influenced Formats (e.g., relying heavily on Auto-suggestions and Predictive Typing). Overall, this visual demonstrates how Platform Dynamics and Technology influence the Development of New English Varieties.

4.3 Qualitative Results

Analyzing the Qualitative Patterns in Relation to Quantitative Outcomes, the Interplay Between the Two Revealed Statistically Significant Patterns of Linguistic Feature Displays Across Social Networking Platforms. Instagram Had the Highest Multimodal Frequency (%) of All Platforms (65%) and Also Had the Largest Concentration of Identity-Indexed Markers, Which Supported a Strong Correlation Between the Multimodality of Digital Communication With Individuals' Stylistic Presentations of

Digital Communication Platforms: Instagram; YouTube; Gaming Chats; and AI Messaging. Instagram and YouTube had the greatest Number of Multimodal expression Types and Lexical Hybrid Types due to

Themselves. YouTube Provided A Slightly Lower Multimodal Frequency (%), However It Had A Significantly Higher Frequency of Evaluative Discourse Structures, Which Correlated With An Average of 40% Lexically Hybridised Words With Referenced Sources Over the Last Month of Data Collection. Qualitative Analyses of Comments on YouTube Index the Following: Predictable Interactional Patterns Support the Predictive Analyses of Hybrid Discourse Forms on YouTube by Statistical Analyses.

Gaming Environments Had the Highest Percentage of Syntactically Reduced Phrase Structure (% 35% Reduction) and Exist as Discourses Where Providers Use High-Speed and Short Forms of Communication (e.g., Acronyms and Clipped Expressions) to Communicate. Patterns of Language in Gaming Environments Stated That the Average Length of a Gaming Communication has an Average of 5.1 Words or Fewer (Statistically). For AI-Enriched Messaging Platforms, analyzed verbatim for this Analysis, Organizations Were Observed as Having a Significant Presence of Algorithmically Derived Expressions by a

Quantitative Analysis, Demonstrating Strong Representation via High Algorithm Signature (% 32%) in the Data. Across Platforms, Qualitative Analyses of the Relationships Among PMSM, Syntactic Efficiency, Hybrid Lexical Forms, and Algorithmic Influence Show That Patterns Exist at a Statistical-Linguistic Level.

4.4 Summary of Findings

In Summary, Current Digital Communication Processes Inflect the Understanding of Language Varieties in English And May Contribute to Creating Varieties of English Through Statistical Patterns Based on User Behaviors. Data Analysis Showed (37.4% Increase) A Substantial Increase in Hybrid Lexical Formations, Resulting from The Commingling of Multilingual User Bases with Patterns of Discourse Within the Digital Communication Space. Correspondingly, the Data Analysis Also Showed a Reduction of Syntactic Complexity of (29.1%) Supported by Statistical Decrease in Length and Subordination of Clauses. The Change Toward Syntactically Minimal Constructs as A Result of The Limitations Imposed By Platforms and User Preferences of Brevity Were Previously Discussed Using Gaming Environment Examples. The Data Develops a Clear Quantifiable Relationship between Digital Communications through Digital "Non-textual" Semiotic Resources. The Influence of Algorithms on Users' Linguistic Patterns Was Confirmed Statistically by Showing Significant Influence (22%) on Retained Phrase Patterns Throughout the Data Collection Period. This Indicates That Users' Use of Predictive Text, Recommending Algorithms, and AI-Generated Content Create a Statistical Imprint on Digital Communication Within English. There Were Significant Statistical Differences Between Platforms for Recorded Patterns Of Syntax. Highest % of Multimodal Density = Instagram, Most Significant % of Syntactic Reduction = Gaming Environment, and Most Significant Algorithmic Signature = AI-Messaging System. Taken All Together, These Data Show That Digital Ecologies, Indeed Create and Are Actively Shaping New English Sub-Varieties Through Users' Quantifiable Behaviors and

Technology Modification.

5. Discussion

5.1 Interpretation of The Statistical Findings

According to the findings of the research, the influence of contemporary digital communications is significant and is having a pronounced impact upon the way that people are writing in English; the findings of this study demonstrate that increased use of hybridized lexical forms accounts for the increased globalization of the English language, i.e. the 37.4% increase in hybridized lexical forms and the increasing use of English language forms combined with those from local languages (the original text states that there has been a substantial increase in Globalization of English through Hybridized Lexical Forms, which is a large increase in the degree to which Globalization of English through Hybridized Lexical Forms will continue to occur), as well as the increase in the extent to which different platform slang is combined with algorithms circulating in the digital space. In combination, the increased use of hybridized forms indicates that the users interacting with these forms are not only multilingual, they are also interacting in a globalized way; therefore, the boundaries between languages are fluid and subject to constant negotiation.

The 29.1% reduction in syntactical complexity reported in the statistical results strongly supports the principle of Digital Efficiency. The reduction indicates that users are using compressed forms and minimal syntactical structures when communicating in fast moving formats such as gaming chats. The reduction demonstrates that the concept of syntactical economy, where users are using fewer words to convey the same amount of meaning, is an essential feature of Digital English Ecosystems.

Another similarly important finding from this research was that of the 22% of recurring phrases that were created by algorithms. This indicates the role of algorithms (e.g. predictive text, recommendation systems, auto-suggest) in creating the linguistic output of users; therefore, we could describe the case of "Technology

shaped Languages” (i.e. the output of the user has been created using a combination of the user's intent and creativity, as well as the input of a machine). The finding that the highest increase in multimodal markers was 56.8% confirmed that we are beginning to see a shift in how individuals are creating meaning (the original text states that Hybrid Semiotic Forms such as Emoji's, Visual cues, and Symbolic Short Hand are continuing to be considered as Integral Components of Digital English). Emerging English varieties emerge as being shaped by user agency, platform culture, and algorithmic facilitation.

5.2 Implications

The statistical and qualitative results indicate that digital sociolinguistic studies have far-reaching implications concerning theories of Sociolinguistics, Language Education, and Technology Development. First, the mechanisms that produce and shape the development of English varieties should be introduced into sociolinguistic models. Digital mediators must be understood as a major player in the mechanisms driving Linguistic Change. Furthermore, traditional views of variation in language based on geography, society or culture are becoming less relevant. Rather, because of the unique nature of digital environments they function as distinct ecological systems with their own norms and rules for emerging, diffusing and stabilizing linguistic diversity in new English varieties. Given how predictive algorithms are able to influence how language is created negative evidence warrants a review of the role those non-human entities (i.e.: algorithms) play in determining the way Linguistic behavior and Discourse Practice will take place.

Second, the results of this study affirm the need for future approaches to language education that are reflective of the realities of contemporary forms of electronic communication. As syntactic reduction is already showing through increased use of multimodal expressions of meaning within the digital arena learners must be instructed in digital literacy that includes understanding of Linguistic Variability as well as Multimodal

Interpretability and critical evaluation of ready-to-wear (algorithm-guided) forms of content. In order to effectively use hybridized forms of communication where the meaning is created through a combination of words, visuals and system-generated prompts, learners will require skills to be able to navigate the challenges presented by these environments.

Finally, the findings provide important considerations for developers of Artificial Intelligence Systems that provide language suggestions to users. Because algorithmic systems provide the content for Language Patterned by Automated Decision Support Systems, developers should adhere to guidelines when designing language tools for the production of automated responses that provide content consistent with regional linguistic diversity, regional dialect, and cultural considerations.

5.3 Contribution to Knowledge

This study represents a substantial contribution to the rapidly emerging field of Digital Sociolinguistics. To the extent that this study provides, for the first time in the Digital Sociolinguistic Literature, a statistically grounded model to explain the collective interactions of Hybridization, Syntactic Reduction, Multimodal Expansion and Algorithmic Influence on the development of New English Varieties across Digital Media platforms, this will serve to increase knowledge of the processes of Language Change, specifically Digital Language Change, and the role of technology in the development and evolution of Language Diversity. Further, the findings of this study will also assist researchers in developing new methodologies and frameworks through which to empirically investigate Linguistic Features and Linguistic Diversity within Technology Mediated Communication. Finally, because of the Cross-Platform Comparative approach used in this study, the researcher was able to develop a means of identifying a number of distinct Digital English Subtypes (e.g. Instagram English, YouTube English, Gaming English, AI-Mediated English) which exist within the unique

Socio-Technological Ecologies of each of the Digital Media Platforms. The presence of supporting Quantitative Data provides empirical evidence for extending the existing Sociolinguistic Scholarship on Digital Language to recognize that Digital Environments are sites of Linguistic Diversification and, therefore, that Digital Environments represent new opportunities for the Diversification of Language, as opposed to the continuation of traditional Language patterns for Off-Line Communication.

6. Conclusion

This research has established a solid basis upon which to evaluate how digital communications today are causing the creation of new forms of the English language. By examining a corpus of 4,800 samples taken from four key social media networks, the project has achieved its goals of discovering new lexicon, syntax, and function in English; investigating how digital communication systems (or algorithms) affect the way people use the English language; analyzing how different types of linguistic modalities (i.e., images, words, sounds) are

combined to create meaning; and analyzing how users on each individual platform behave linguistically. The statistical data are consistent with the claims made in the study, namely, that digital English is evolving based upon quantifiable relationships between users of digital communication networks and technology. The findings indicate that linguistic globalization, digital efficiency, and algorithmic mediation drive language change. Furthermore, the cross-platform comparison shows that Instagram, YouTube, gaming platforms and AI-based messaging systems represent distinct sociolinguistic ecologies that offer users opportunities for innovation and diversification in English. Thus, this study is the first to provide a statistically supported model for the ways in which technology maintains an ongoing real-time cycle of innovation for the English language. This research also informs sociolinguistic theory and has ramifications for the development of linguistic inclusive or accessible Artificial Intelligence systems. Furthermore, the findings illustrate how digital communications are creating new types of English.

References

- Tagg, C., & Evans, M. (2021). Everyday language in mobile messaging: A corpus-based study. *Mobile Media & Communication*, 9(1), 52–71. <https://doi.org/10.1177/2050157920949954>
- Androutsopoulos, J. (2020). Online data collection as sociolinguistic practice. *Journal of Sociolinguistics*, 24(5), 555–573. <https://doi.org/10.1111/josl.12383>
- Herring, S. C., & Dainas, A. (2020). Language, interaction, and digital context. *Annual Review of Linguistics*, 6, 331–350. <https://doi.org/10.1146/annurev-linguistics-011619-030542>
- Sergeant, P. (2020). The influence of social media communication on contemporary English. *World Englishes*, 39(2), 267–282. <https://doi.org/10.1111/weng.12432>
- Leppänen, S., & Kytölä, S. (2020). Digital ethnography and English language variation. In *The Routledge Handbook of English Language Studies* (pp. 290–305). Routledge. <https://doi.org/10.4324/9780429341573-23>
- Takahashi, K. (2021). English in digital translingual practices on social media. *Applied Linguistics Review*, 12(4), 477–499. <https://doi.org/10.1515/applirev-2020-0051>
- Georgakopoulou, A. (2021). Small stories on Instagram and identity construction. *Social Media + Society*, 7(1), 1–11. <https://doi.org/10.1177/2056305121988933>
- Heyd, T. (2020). Digital Englishes and linguistic innovation online. *English World-Wide*, 41(1), 1–24. <https://doi.org/10.1075/eww.00043.hey>
- Page, R. (2021). Narratives and identity in online English varieties. *Discourse, Context & Media*, 42, Article 100517. <https://doi.org/10.1016/j.dcm.2021.100517>
- Squires, L. (2020). English variation across digital platforms. *Annual Review of Applied Linguistics*, 40, 160–179.

<https://doi.org/10.1017/S0267190520000063>

Eisenstein, J. (2019). Social media and language variation: Large-scale patterns. *Annual Review of Linguistics*, 5, 26–45. <https://doi.org/10.1146/annurev-linguistics-011718-012542>

Blommaert, J. (2020). Digital repertoires and algorithmic communication. *Tilburg Papers in Culture Studies*, 234. <https://doi.org/10.31235/osf.io/2b8qv>

Varis, P. (2021). Algorithmic cultures and online English. *Social Semiotics*, 31(3), 453–469. <https://doi.org/10.1080/10350330.2020.1766190>

Bolton, K., & Botha, W. (2020). The emergence of new English forms in digital Asia. *Asian Englishes*, 22(2), 87–106. <https://doi.org/10.1080/13488678.2019.1696133>

Lee, C. (2021). Multimodal discourse and online English mixing. *Discourse Studies*, 23(3), 343–361. <https://doi.org/10.1177/1461445620972658>