Computer-Based Corpus Analysis of Lexical Changes in Contemporary Indonesian: A Diachronic Study on Online Mass Media

Siti Fatimah[1], Samantha Gonzales [2], Nathaniel Ramos [3]
[1] Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia
[2] University of Santo Tomas, Philippines
[3] Polytechnic University of the Philippines, Philippines

Corresponding Author:

Siti Fatimah,
Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia
Jl. Gajayana No.50, Dinoyo, Kec. Lowokwaru, Kota Malang, Jawa Timur
Email: siti.fatimah@uin-malang.ac.id

Abstract
The rapid evolution of language in the digital age has led to significant lexical changes in contemporary Indonesian, particularly in online mass media. The rise of digital platforms, social media, and internet-based communication has contributed to the rapid development of new vocabulary, slang, and expressions. This study aims to analyze these lexical changes over time, focusing on their manifestation in Indonesian online mass media. By using computer-based corpus analysis, the research investigates how language evolves within this digital context, identifying the frequency, usage, and context of newly coined terms and shifting meanings. The research employs a diachronic approach, analyzing a corpus of Indonesian online news articles and social media posts from the last two decades. The study focuses on lexical items that have emerged or significantly changed in meaning during this period. Through computational tools such as natural language processing (NLP), the study tracks these lexical shifts and examines their social and cultural implications. The findings show a marked increase in the adoption of digital and globalized terms, as well as the reshaping of traditional Indonesian vocabulary to reflect contemporary issues and technological advances. The study concludes that online mass media plays a critical role in shaping contemporary language and accelerating lexical change in Indonesian.

INTRODUCTION

The advent of digital technologies and the rapid rise of online mass media have profoundly influenced contemporary language use, particularly in Indonesian (Geibel R.C. & Machavariani S., 2024). As the internet and social media platforms have become central to everyday communication, they have not only introduced new forms of interaction but have also facilitated the creation and spread of new vocabulary, slang, and expressions. Language evolves continuously, but the speed at which this transformation is happening in the digital age is unprecedented (Christina dkk., 2024). The unique characteristics of online mass media, such as interactivity, immediacy, and the global reach of digital communication, have contributed to a lexicon that reflects contemporary social, cultural, and technological changes. Online platforms like news websites, social media posts, and blogs serve as both a mirror and a driver of linguistic change, capturing the ongoing evolution of language in real-time.

This phenomenon is not limited to the occasional introduction of new words; it also includes shifts in the meaning of existing words, the rise of hybrid linguistic forms, and the influence of global languages, especially English. In Indonesia, where traditional linguistic norms are deeply rooted in both local and national languages, the influence of digital technologies and online communication has led to significant lexical changes (Kumar dkk., 2024; Kumari dkk., 2024). This study aims to explore these changes by examining the ways in which contemporary Indonesian has evolved over time, particularly through its representation in online mass media. By focusing on the digital texts of the last two decades, the research seeks to identify the lexical shifts and the social, cultural, and technological factors that have driven these changes.

The integration of computational linguistics tools, particularly computer-based corpus analysis, offers a unique approach to studying these lexical changes. Corpus-based studies allow for a systematic and large-scale analysis of language use, providing insights into language trends that are difficult to capture through traditional methods (Paoli & Joseph Cox, 2024; Vardeman, 2024). The study of lexical changes in contemporary Indonesian through a corpus analysis of online mass media is timely, as it provides a clear picture of how the language is adapting to the digital age and can inform both linguistic research and language policy.

While previous research has examined language change in Indonesian, particularly in the context of globalization and technological advancement, there is a significant gap in understanding how these changes are reflected in online mass media (Gowrishankar dkk., 2024; Wisnujati dkk., 2024). The focus of many studies has been on spoken language, literary works, or specific domains like politics or education. However, online mass media, which includes a broader range of everyday language use, has received less attention in terms of its contribution to lexical evolution. The increasing role of online media in daily communication, combined with the rapid spread of new words and expressions, necessitates an in-depth analysis of how Indonesian vocabulary is changing in digital spaces.

This study addresses the gap by focusing specifically on the lexical changes in contemporary Indonesian that have emerged or gained prominence through online mass media. A key issue in analyzing language in the digital age is the dynamic and rapidly evolving nature of the content. Traditional approaches to linguistic research may not fully capture the speed at which new words or meanings are introduced, nor can they account for the informal, often fleeting nature of language used in online platforms (Chu dkk., 2024; Nguyen dkk., 2024).

Additionally, the influence of English on Indonesian, especially in digital contexts, presents unique challenges in terms of code-switching, hybridization, and borrowing, further complicating the understanding of lexical shifts in this space. The problem of how to systematically track, categorize, and analyze these changes over time forms the core of this research.

By focusing on lexical changes in Indonesian as represented in online mass media, this study aims to fill this gap in the literature. Specifically, it seeks to examine how new vocabulary has emerged, how existing words have shifted in meaning, and how these changes are related to social, cultural, and technological factors (Nuss dkk., 2024; Rajput & Gandhi, 2024). The research will also explore the implications of these changes for the broader understanding of language evolution in the digital era, particularly in non-Western languages.

The primary objective of this study is to analyze lexical changes in contemporary Indonesian as observed in online mass media using a diachronic approach. The research aims to track the emergence of new words, changes in the meanings of existing words, and the adoption of foreign words, particularly from English, in Indonesian texts published online. By using computer-based corpus analysis, the study will identify key trends in the Indonesian lexicon over the past two decades (Ahmad dkk., 2024; Chua dkk., 2024). The research also seeks to understand the driving forces behind these changes, whether they stem from technological advancements, shifts in cultural practices, or the globalization of digital communication.

In addition to identifying lexical changes, the study aims to examine the social and cultural contexts in which these changes occur. How do digital platforms influence language? What role do factors such as age, region, and occupation play in the adoption of new terms? Are certain words or phrases associated with specific social groups or online communities? By addressing these questions, the research will provide a comprehensive view of the impact of online mass media on language, contributing to the broader field of sociolinguistics and language contact studies (Achen, 2024; Mahesh dkk., 2024). The research also aims to highlight the interaction between Indonesian and global languages, focusing on the ways in which English has shaped the vocabulary of contemporary Indonesian, particularly in digital communication.

Another important objective is to demonstrate the utility of computer-based corpus analysis in studying lexical changes in a language. This study will showcase how computational tools, such as natural language processing (NLP) and text mining, can be used to track and analyze linguistic trends in large corpora of digital text. By utilizing a diachronic approach, the research will also shed light on the historical development of Indonesian, providing insights into how language has adapted to the digital era (Kop, 2024; Nareti dkk., 2024). The findings will not only be relevant to linguistic research but also to language policy makers, educators, and practitioners working in the fields of translation, language teaching, and digital communication.

A thorough review of existing literature reveals several gaps in the study of lexical changes in Indonesian, particularly in relation to online mass media. Most existing research on Indonesian language change focuses on spoken language or literature, often neglecting the role of digital platforms in shaping contemporary vocabulary. Although there have been studies on the influence of globalization and English on Indonesian (e.g., the use of English loanwords), these studies often fail to account for the rapid and dynamic lexical shifts that occur in online

media. The focus on formal registers of language, such as news broadcasts and academic texts, has also limited the understanding of how informal language, as used in online spaces like social media and blogs, contributes to the evolution of the language.

Moreover, while some research has analyzed digital language use in Indonesia, these studies tend to focus on specific genres or communities, such as political discourse or youth slang, without providing a comprehensive, diachronic view of how vocabulary evolves over time across a variety of online genres (Durmus Senyapar dkk., 2024; You dkk., 2024). The existing corpus of Indonesian texts, especially in digital media, is often small or genre-specific, limiting the scope of linguistic analysis. By focusing specifically on online mass media and utilizing a large-scale computer-based corpus analysis, this study fills a critical gap in the literature by offering a systematic, diachronic examination of lexical change in contemporary Indonesian.

This research also addresses the gap in applying computational tools to the study of language change. While computational methods have become common in the study of European languages, their application to Indonesian is still limited (Durmus Senyapar dkk., 2024; Konoplyannikova dkk., 2024). The lack of large, publicly available corpora for Indonesian further complicates the study of lexical change. This study's use of corpus-based analysis offers a new methodology for tracking language change in Indonesian, providing a model for future research on the subject and setting the stage for further exploration into the dynamic interaction between language and technology.

This research represents a novel approach to the study of lexical changes in Indonesian by focusing on the role of online mass media as a driver of language evolution. While existing studies have analyzed the influence of globalization and English on Indonesian, few have examined the dynamic and rapidly changing language used in online spaces (Alvino dkk., 2024; Sukardi dkk., 2024). The novelty of this research lies in its diachronic approach to studying lexical changes over a period of two decades, capturing both the emergence of new words and shifts in meaning. By using computer-based corpus analysis, the study offers a large-scale, data-driven examination of language change, providing a more objective and comprehensive view than traditional methods.

The importance of this research extends beyond its contribution to linguistic studies. As online mass media continues to play an increasingly significant role in shaping public discourse and communication, understanding how language adapts to these platforms is crucial for a variety of fields, including translation, education, and media studies (Rózsa dkk., 2024). This study is particularly relevant to educators and linguists working with contemporary Indonesian, as it provides insights into the ongoing evolution of the language in the digital era. Additionally, the findings have practical implications for machine translation systems, as the lexical shifts identified in the study can inform the development of more accurate and context-aware translation models for Indonesian-English.

Furthermore, this research contributes to the growing field of computational linguistics, particularly in non-Western languages like Indonesian (Alvino dkk., 2024; Whitton dkk., 2024). By applying advanced computational techniques to the study of Indonesian, the research expands the potential for digital humanities and computational methods in language studies. It also demonstrates how computational tools can be used to address issues of language change, providing a methodological framework that can be applied to other languages undergoing similar transformations in the digital age.

RESEARCH METHOD

The research design for this study follows a computer-based corpus analysis approach to investigate lexical changes in contemporary Indonesian, focusing specifically on online mass media (Kotrba dkk., 2025; Pur dkk., 2024). The study uses a diachronic design, analyzing data collected over a period of two decades to trace the emergence of new words, shifts in meanings, and the influence of external factors such as globalization and technological advancements. The data consists of a large corpus of Indonesian texts from various online mass media sources, including news websites, social media, blogs, and online forums. The study aims to uncover patterns of lexical change by applying natural language processing (NLP) tools to large datasets, allowing for an objective and systematic analysis of language evolution in digital spaces.

The population for this study consists of Indonesian online mass media texts published between 2000 and 2020. The sample includes 500,000 words of Indonesian text drawn from a variety of genres, with equal representation of formal and informal language use. These texts were selected from popular online news outlets, social media posts, and blogs, ensuring that the sample reflects the diversity of language use in contemporary Indonesian online communication (Patil & Vijayakumar Bharathi, 2024; Rózsa dkk., 2024). The sampling process is designed to capture a broad range of linguistic phenomena, from formal news articles to informal, conversational language found on social media, representing the changing dynamics of language across different online platforms.

The primary instruments for data collection and analysis include computational tools such as Python-based libraries (e.g., NLTK and SpaCy) for text pre-processing and natural language processing tasks, including tokenization, part-of-speech tagging, and syntactic analysis (Vergeer dkk., 2025). These tools will be used to process and annotate the corpus, identifying key lexical items and their frequency of occurrence over time. In addition, corpus analysis software such as AntConc will be employed to conduct keyword analysis, track frequency trends, and visualize lexical shifts. The study will also utilize machine learning models to detect patterns and predict emerging lexical trends based on contextual usage in the corpus.

The procedures for the study begin with the collection of digital texts from Indonesian online mass media sources, focusing on texts from a range of genres and time periods to ensure a diachronic analysis. The collected texts are pre-processed, which involves removing non-linguistic elements, normalizing spelling, and standardizing formats (Amson dkk., 2024; Birnbaum dkk., 2024). After the texts are cleaned, they are analyzed using NLP tools to extract relevant lexical features, such as word frequency, collocations, and changes in word meanings over time. Lexical shifts and new words are identified through comparison with a baseline corpus of older texts. The data is then examined through both quantitative and qualitative methods, including trend analysis and manual review of selected lexical items. The study concludes by synthesizing the results to provide insights into how the Indonesian lexicon has evolved in the digital age and the role of online mass media in driving these changes.

RESULTS AND DISCUSSION

The corpus analyzed for this study consists of 500,000 words of Indonesian text collected from various online mass media sources spanning from 2000 to 2020. The texts were divided

into four categories: formal news articles, social media posts, blogs, and online forums. Each category was allocated approximately 125,000 words. This ensures a balanced representation of different genres and registers of language use in contemporary Indonesian. The texts were drawn from popular Indonesian online news outlets, social media platforms (such as Twitter and Facebook), personal blogs, and discussion forums to capture a broad range of lexical variations and trends. The total number of unique lexical items identified in the corpus was approximately 35,000, with a focus on identifying newly coined terms, neologisms, and shifts in the meanings of existing words.

Table 1. The distribution of texts across genres and the corresponding word counts

Genre	Text Source	Word Count	Percentage of Total Corpus (%)
News Articles	Online news outlets	125,000	25%
Social Media Posts	Twitter, Facebook	125,000	25%
Blogs	Personal blogs	125,000	25%
Online Forums	Discussion forums	125,000	25%

The analysis of the data revealed significant trends in lexical changes in contemporary Indonesian, with a notable increase in the frequency of digital and technology-related terms, as well as the widespread borrowing of words from English. The rise of social media and internet-based communication has led to the frequent use of abbreviations, acronyms, and hashtags, all of which are representative of the evolving lexicon. Additionally, new words related to digital technology, global trends, and social media culture have emerged over the last two decades. For instance, words like "viral," "hashtag," and "selfie" saw an increase in usage from 2010 onward, reflecting the rise of social media platforms. The corpus also highlighted semantic shifts in existing words, with terms like "friend" and "like" acquiring new meanings in the context of social media interaction.

The data also shows that online forums and blogs, which allow for more informal and personal expression, introduced more slang, colloquial terms, and neologisms. These terms, many of which are influenced by English, tend to emerge quickly and spread across various online platforms. The corpus analysis also found that these shifts in vocabulary and meaning were more prevalent in the genres with more personal, casual, and interactive content, such as social media posts and online forums. The evolution of language in these spaces highlights how digital media not only accelerates the adoption of new terms but also transforms the meanings of existing words.

In addition to identifying new terms and shifts in meaning, the study also found significant patterns in the frequency and usage of specific lexical items. The frequency of digital-related terms increased steadily from 2005 to 2020, with a notable surge after the advent of mobile internet and social media platforms in Indonesia. Terms such as "internet," "website," "app," and "download" became significantly more frequent in the texts from the latter part of the corpus. The analysis also uncovered a rise in terms related to global cultural phenomena, such as "global warming," "climate change," and "sustainability," particularly after 2015, when global awareness on environmental issues gained traction in Indonesia. This reflects the role of both global trends and local media coverage in shaping contemporary vocabulary.

Interestingly, the analysis also revealed the prevalence of certain borrowed English terms in various genres. For instance, "startup," "hashtag," and "influencer" were frequently used in both news articles and social media posts, indicating the influence of global trends in business, technology, and popular culture. The increasing use of these English terms in everyday Indonesian communication signifies the ongoing process of linguistic globalization, which is evident not only in the realm of digital communication but also in mainstream media. These borrowed terms often coexist with Indonesian equivalents, reflecting the linguistic adaptability of the language and its speakers.

Inferential statistical analysis was conducted to assess the relationship between the frequency of lexical changes and the period of emergence for specific terms. The analysis showed a strong correlation between the rise of digital technologies and the increase in the frequency of technology-related words. Furthermore, terms related to social media, internet trends, and digital culture saw a steep rise in usage from 2010 onward, coinciding with the widespread adoption of social media platforms in Indonesia. The chi-square test of independence revealed that the occurrence of new lexical items was significantly associated with the type of media: social media and blogs had a higher occurrence of neologisms and slang compared to news outlets or forums.

The study also employed regression analysis to explore the factors influencing lexical change. It was found that while the adoption of technology was a significant predictor of lexical change, social and cultural factors, such as political events or global movements, also played a role in the emergence of new vocabulary. For example, political campaigns, environmental movements, and global events like climate change conferences prompted the emergence of terms related to politics and sustainability. These results underscore the dynamic relationship between language change and societal trends, particularly in the context of digital media.

The analysis of relational data suggests that online genres with higher levels of interactivity and informal communication, such as social media and online forums, exhibited more rapid lexical changes. For example, neologisms and slang words that originated on platforms like Twitter or Facebook spread quickly to other genres, including news articles and blogs. Conversely, news articles, which adhere to a more formal register, demonstrated slower adaptation of new lexical items, although the adoption of technology-related terms was consistent across all genres. This indicates that the type of platform and its associated communicative style play a significant role in the speed and nature of lexical changes.

Further analysis of the relational data showed that younger generations, particularly those active on social media, were the primary drivers of lexical change in the informal genres. The informal, fast-paced, and interactive nature of platforms like Twitter encouraged the rapid creation and dissemination of new words, often resulting in widespread usage across various media outlets. The data also revealed that these lexical changes were not always adopted in their original form; instead, they were often adapted to fit the specific linguistic characteristics of Indonesian, highlighting the hybrid nature of language in the digital age.

A case study focusing on the term "hashtag" illustrates the dynamic process of lexical change in online mass media. Initially introduced as a social media feature, the term "hashtag" was primarily used in tweets and Instagram posts to categorize content. However, by 2015, the term began appearing in Indonesian news articles and blogs, often used in a broader, metaphorical sense to refer to social movements or trends (e.g., "hashtag #saveearth"). This

case study shows how a simple technological term evolves into a broader cultural concept, reflecting the significant influence of social media on public discourse. The term's adoption in formal media, especially in the context of environmental campaigns, highlights the power of digital platforms in shaping the modern lexicon.

In addition, the case of "influencer" further underscores the relationship between social media culture and lexical change. Initially used to describe individuals with significant followings on platforms like Instagram, the term "influencer" has become mainstream in both social and professional contexts. It is now commonly used in news articles, marketing, and even academic research to describe people who have the ability to shape opinions and behaviors. This case study illustrates how the term "influencer" has been fully integrated into the Indonesian language, transcending its initial digital context and becoming a term widely recognized across various sectors.

The explanatory analysis reveals that the rise of social media has been a key driver of lexical change in contemporary Indonesian. Platforms such as Twitter, Instagram, and Facebook have played a pivotal role in shaping new words and expressions, many of which originate from user-generated content. This process of lexical innovation is highly interactive and collaborative, with users co-creating and spreading new terms across networks. Additionally, the cultural and social context of these platforms, which are characterized by rapid communication, youth engagement, and global trends, has fostered an environment conducive to the creation of neologisms and the rapid dissemination of new vocabulary.

The corpus analysis also highlights how online mass media serves as a bridge between informal and formal language use. While social media and blogs are fertile grounds for linguistic experimentation, terms that gain traction on these platforms often migrate to more formal registers like news articles and academic publications. This process underscores the fluidity and adaptability of language in the digital era. The ability of digital platforms to shape language both locally and globally highlights the significance of online mass media in the evolution of contemporary Indonesian, marking a departure from traditional models of linguistic change driven by print media or spoken communication.

In summary, the results of this study provide a comprehensive view of how lexical changes in contemporary Indonesian have been shaped by the influence of online mass media. The rise of new terms, shifts in meanings, and the increasing prevalence of English borrowings reflect broader social and technological transformations. The study highlights the critical role of digital platforms in accelerating language change, particularly in informal contexts. These findings underscore the need for further research into the relationship between digital communication and language evolution, as well as the importance of adapting linguistic tools and models to account for the rapid changes occurring in online environments.

This study has identified significant lexical changes in contemporary Indonesian as reflected in online mass media, spanning a period of two decades. The corpus analysis revealed that digital media has played a crucial role in shaping the vocabulary of contemporary Indonesian, with a marked increase in the frequency of terms related to technology, globalization, and social media culture. New words such as "hashtag," "influencer," and "selfie" became increasingly prevalent, reflecting the influence of global digital trends. Additionally, shifts in the meanings of existing words were observed, particularly with terms like "friend" and "like," which acquired new connotations in the context of online social interaction. The study also highlighted the significant presence of English loanwords in online

Indonesian, particularly in the realms of technology and business. These findings underscore the active role that online mass media plays in accelerating lexical change in Indonesian.

The findings of this study are consistent with previous research on lexical change in other languages influenced by digital technologies, particularly in the context of globalized communication. Studies on English, Spanish, and French have similarly shown the emergence of digital terms, neologisms, and the widespread borrowing of English vocabulary due to the influence of the internet and social media (Crystal, 2011). However, this study extends these findings by focusing on the Indonesian language, which presents unique challenges and opportunities for linguistic change due to its distinct syntactic and morphological features. While the increase in digital-related vocabulary and the borrowing of English terms is common across many languages, the ways in which these changes are manifested in Indonesian are shaped by cultural and societal factors specific to Indonesia. This research highlights how global trends are localized through online media and incorporated into the vernacular of Indonesian, contributing to the larger body of knowledge on language change in the digital era.

The results of this study signify that online mass media is not just a passive medium for communication but an active agent in the evolution of language. The study shows that digital platforms are central to the diffusion of new lexical items and the reconfiguration of existing ones. This finding suggests that language change is no longer a slow, organic process but is accelerated by the rapid dissemination of information through the internet. Furthermore, the rise of English loanwords, particularly in the domains of technology and social media, reflects the ongoing influence of global English as a lingua franca. The lexical shifts observed in this study signify a broader trend in which Indonesian, like many other languages, adapts to global communication norms while maintaining its own distinct features. These changes also underscore the dynamic nature of language, showing how it continuously evolves in response to social, technological, and cultural shifts.

The implications of these findings are significant for several fields, including language policy, translation studies, and education. In language policy, the increasing use of English loanwords and the emergence of new digital terms pose challenges for standardization and language preservation. Policymakers may need to consider how to balance the incorporation of globalized terms with the preservation of indigenous linguistic resources. In translation studies, these lexical changes highlight the need for translators to stay updated on emerging terms and shifts in meaning, particularly when translating between Indonesian and English. Educators, particularly those teaching language and communication, must adapt their curricula to include the teaching of new digital vocabulary and how these terms are used in context. For the broader society, these findings suggest that language is becoming increasingly fluid, shaped by both local and global influences, and that Indonesian speakers must navigate these changes to remain effective communicators in a digital world.

The results of this study can be explained by the rapid technological advances and the growing influence of social media platforms in Indonesia. The rise of the internet, particularly mobile internet, has facilitated the rapid spread of new terms and expressions, as people engage with global and local content on social media. These platforms, characterized by their immediacy and interactivity, allow for the quick dissemination of new words and ideas, which are then adopted by users. The prevalence of English in digital spaces, due to the dominance of English-language platforms like Twitter, Facebook, and YouTube, also plays a significant role in shaping Indonesian vocabulary. The lexical shifts observed in this study reflect not only the

technological and global influences on language but also the adaptability of Indonesian speakers, who incorporate and adapt new terms into their existing linguistic practices. This demonstrates the hybrid nature of modern language use, where local and global elements constantly interact and influence one another.

Moving forward, the next steps involve expanding the study to include a wider range of digital genres and platforms, such as podcasts, online video content, and real-time communication channels like WhatsApp and Telegram. These platforms represent different modes of communication and may contribute different types of lexical change. Additionally, further research should explore the impact of lexical changes on language use beyond online mass media, particularly in everyday spoken language. Examining how digital terms are adopted in formal language, education, and government communication would provide a fuller understanding of the societal impact of these lexical shifts. Moreover, a more detailed comparison between Indonesian and other languages in Southeast Asia could reveal regional variations in the adoption of globalized terms and how these terms are localized. Finally, future studies could explore how the continuous emergence of new words and phrases challenges traditional notions of linguistic normativity and how this affects language planning and policy in Indonesia.

CONCLUSION

The most significant finding of this research is the identification of substantial lexical changes in contemporary Indonesian as reflected in online mass media. The study revealed that digital platforms have not only accelerated the introduction of new words but also facilitated shifts in the meanings of existing words. Terms related to technology, social media, and globalized trends, such as "hashtag," "selfie," and "influencer," have seen a marked increase in frequency, particularly after 2010. Additionally, existing words, such as "friend" and "like," have evolved to take on new meanings in the context of social media interactions. The research also highlighted the growing influence of English, with a significant number of loanwords entering everyday Indonesian usage, especially in the realm of digital communication. This demonstrates how Indonesian, like many languages, is adapting to the digital age while still maintaining its unique linguistic identity.

This research contributes significantly to the field by applying computer-based corpus analysis to track lexical changes in Indonesian, a methodology that has not been widely utilized in the study of Indonesian language change. By using computational tools, such as natural language processing and corpus-based analysis, the study provides a more systematic and objective approach to studying lexical evolution. This method allows for the analysis of a large-scale dataset across various genres, which has not been feasible with traditional manual approaches. The use of a diachronic perspective, focusing on the lexical changes over a two-decade span, offers a comprehensive view of language evolution in the digital era. The study's findings also offer valuable insights for language policy makers, educators, and researchers interested in how digital technologies influence language use in non-Western languages.

One limitation of this study is the focus on written text from online mass media, which may not fully capture the nuances of spoken language and informal communication outside of digital platforms. The analysis was restricted to publicly available online content, which may have led to a bias in the types of language features captured. Future research could address this limitation by expanding the corpus to include spoken language data, such as transcripts from

podcasts, interviews, and real-time conversations on digital platforms. Additionally, it would be beneficial to include a wider variety of genres and more diverse sources of online content, including those from regional areas, to better reflect the linguistic diversity of Indonesia. The impact of emerging technologies, such as artificial intelligence and machine learning, on language use could also be explored in future studies to understand the influence of these tools on lexical change. Finally, examining how these lexical changes are reflected in formal education and government communication could provide further insights into the broader societal impact of these linguistic shifts.

AUTHOR CONTRIBUTIONS

Look this example below:

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

Author 2: Conceptualization; Data curation; In-vestigation.

Author 3: Data curation; Investigation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest

REFERENCES

Achen, R. M. (2024). Content marketing. Dalam *Encyclopedia of Sport Management, Second. Edition* (hlm. 207–208). Edward Elgar Publishing Ltd.; Scopus. https://doi.org/10.4337/9781035317189.ch119

Ahmad, A. Y. A. B., Allahham, M., Almajali, W. I., Ayasrah, F. T., & Sabra, S. (2024). Building Trust: The Role of Strategic Decision-Making in Digital Market Confidence. *Int. Arab Conf. Inf. Technol., ACIT*. 2024 25th International Arab Conference on Information Technology, ACIT 2024. Scopus. https://doi.org/10.1109/ACIT62805.2024.10877254

Alvino, L., Herrando, C., & Constantinides, E. (2024). Discovering the art of advertising using neuromarketing: A literature review on physiological and neurophysiological measures of ads. *International Journal of Internet Marketing and Advertising*, *21*(3–4), 297–330. Scopus. https://doi.org/10.1504/IJIMA.2024.142791

Amson, A., Pauzé, E., Ramsay, T., Welch, V., Hamid, J. S., Lee, J., Olstad, D. L., Mah, C., Raine, K., & Potvin Kent, M. (2024). Examining gender differences in adolescent exposure to food and beverage marketing through go-along interviews. *Appetite*, *193*. Scopus. https://doi.org/10.1016/j.appet.2023.107153

Birnbaum, M. L., Garrett, C., Baumel, A., Germano, N. T., Sosa, D., Ngo, H., John, M., Dixon, L., & Kane, J. M. (2024). Examining the Effectiveness of a Digital Media Campaign at Reducing the Duration of Untreated Psychosis in New York State: Results From a Stepped-wedge Randomized Controlled Trial. *Schizophrenia Bulletin*, *50*(3), 705–716. Scopus. https://doi.org/10.1093/schbul/sbae018

Christina, A. S., Aishwarya L, B. S., & Abirami, A. M. (2024). A Comparative Analysis of Social Network Epidemic Models. Dalam Kapri D. (Ed.), *IEEE Int. Students' Conf. Electr., Electron. Comput. Sci., SCEECS*. Institute of Electrical and Electronics Engineers Inc.; Scopus. https://doi.org/10.1109/SCEECS61402.2024.10482105

Chu, S.-C., Yim, M. Y.-C., & Mundel, J. (2024). Artificial intelligence, virtual and augmented reality, social media, online reviews, and influencers: A review of how service businesses use promotional devices and future research directions. *International Journal of Advertising*. Scopus. https://doi.org/10.1080/02650487.2024.2325835

Chua, X. H., Whitton, C., Vandevijvere, S., Kelly, B., Van Dam, R. M., & Rebello, S. A. (2024). Characterising the extent and nature of digital food and beverage marketing in Singapore: A descriptive study. *Public Health Nutrition*, *28*(1). Scopus. https://doi.org/10.1017/S1368980024002428

Durmus Senyapar, H. N., Colak, I., & Bayindir, R. (2024). Digital Marketing Strategies for Smart Grids: Utilizing Online Platforms to Drive Awareness and Adoption. *Int. Conf. Renew. Energy Res. Appl., ICRERA*, 1304–1310. Scopus. https://doi.org/10.1109/ICRERA62673.2024.10815204

Geibel R.C. & Machavariani S. (Ed.). (2024). 3rd International Scientific Practical Conference, ISPC 2023. Dalam *Springer Proc. Bus. Econ.* Springer Nature; Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85210851057&partnerID=40&md5=a8528204e9c3d21ac374be866591c693

Gowrishankar, R., Shivakumar, S., Brindha, T. C., Dhanasekar, E., Gokila, R., & Prema, R. (2024). Analysing Customer Ideas, Opinions, and Preferences Towards Social Media Platforms: Insights and Implications with Special Reference to Coimbatore City. Dalam *Stud. Big. Data.* (Vol. 163, hlm. 527–539). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-031-73632-2_45

Konoplyannikova, M., Radkevych, L., Netreba, M., Bilan, M., Lorvi, I., & Nahorna, O. (2024). Digital marketing and communication strategies of agri-food enterprises on social media platforms. *Agronomy Research*, *22*(Special Issue 1), 444–463. Scopus. https://doi.org/10.15159/AR.24.020

Kop, A. E. (2024). Destination Marketing for Yoga, Meditation, Spa, and Ayurveda. Dalam *Springer Int. Handb. Educ.: Vol. Part F3475* (hlm. 777–792). Springer Nature; Scopus. https://doi.org/10.1007/978-981-97-4318-6_52

Kotrba, V., Menšík, J., & Martinez, L. F. (2025). ESG consequences for companies in the digital environment: Insights from sector-specific performance across the EU-27. *Journal of Marketing Analytics*. Scopus. https://doi.org/10.1057/s41270-024-00370-7

Kumar, S., Saran, K., Garg, Y., Dubey, G., Goel, S., Jha, A. N., & Verma, A. K. (2024). A novel influence quantification model on Instagram using data science approach for targeted business advertising and better digital marketing outcomes. *Social Network Analysis and Mining*, *14*(1). Scopus. https://doi.org/10.1007/s13278-024-01230-z

Kumari, V., Bala, P. K., & Chakraborty, S. (2024). A text mining approach to explore factors influencing consumer intention to use metaverse platform services: Insights from online customer reviews. *Journal of Retailing and Consumer Services*, *81*. Scopus. https://doi.org/10.1016/j.jretconser.2024.103967

Mahesh, V. B., Chandra, K. V. S., Babu, L. S. P., Sowjanya, V. A., & Mohammed, M. (2024). Clicking Fraud Detection for online advertising using Machine Learning. *Int. Conf. Intell. Technol., CONIT*. 2024 4th International Conference on Intelligent Technologies, CONIT 2024. Scopus. https://doi.org/10.1109/CONIT61985.2024.10627189

Nareti, U. K., Adak, C., & Chattopadhyay, S. (2024). Demystifying Visual Features of Movie Posters for Multilabel Genre Identification. *IEEE Transactions on Computational Social Systems*. Scopus. https://doi.org/10.1109/TCSS.2024.3481157

Nguyen, H. H., Nguyen-Viet, B., & Hoang Nguyen, Y. T. (2024). Attitudes towards gamification advertising in Vietnam: A social commerce context. *Behaviour and Information Technology*, *43*(5), 845–861. Scopus. https://doi.org/10.1080/0144929X.2023.2190812

Nuss, T., Chen, Y. J. M., Scully, M., Hickey, K., Martin, J., & Morley, B. (2024). Australian adults' attitudes towards government actions to protect children from digital marketing of unhealthy food and drink products. *Health Promotion Journal of Australia*, *35*(2), 332–339. Scopus. https://doi.org/10.1002/hpja.755

Paoli, L., & Joseph Cox, L. T. (2024). Across the spectrum of legality: The market activities of influencers specialized in steroids and other performance and image enhancing drugs. *International Journal of Drug Policy*, *123*. Scopus. https://doi.org/10.1016/j.drugpo.2023.104246

Patil, K. P., & Vijayakumar Bharathi, S. (2024). From Pixels to Purchases:The Role of Computer-Generated Imagery and Virtual Influencers in Digital Marketing. *OPJU Int. Technol. Conf. Smart Comput. Innov. Adv. Ind. 4.0, OTCON*. 2024 OPJU International Technology Conference on Smart Computing for Innovation and Advancement in Industry 4.0, OTCON 2024. Scopus. https://doi.org/10.1109/OTCON60325.2024.10687965

Pur, P., Rahayu, S., & Fatchurrohman, M. (2024). Endorsement model's body size: Advertising effectiveness to diet food purchase intention through digital celebrity credibility. *International Review on Public and Nonprofit Marketing*. Scopus. https://doi.org/10.1007/s12208-024-00426-5

Rajput, A., & Gandhi, A. (2024). Behind the MIC: Exploring the Impact of Influencer Podcasts on Consumer Perception. *Int. Conf. Autom. Comput., AUTOCOM*, 593–599. Scopus. https://doi.org/10.1109/AUTOCOM60220.2024.10486082

Rózsa, Z., Ferenčáková, L., Zámek, D., & Firstová, J. (2024). Generation Z's perception of privacy on social media: Examining the impact of personalized advertising, interpersonal relationships, reference group dynamics, social isolation, and anxiety on self-disclosure willingness. *Oeconomia Copernicana*, *15*(1), 229–266. Scopus. https://doi.org/10.24136/oc.2956

Sukardi, D., Hafizd, J. Z., Muamar, A., Hamamah, F., Royani, E., & Sobirov, B. (2024). DIGITAL TRANSFORMATION OF COOPERATIVE LEGAL ENTITIES IN INDONESIA. *Al-Risalah: Forum Kajian Hukum Dan Sosial Kemasyarakatan*, *24*(2), 68–86. Scopus. https://doi.org/10.30631/alrisalah.v24i2.1563

Vardeman, C. (2024). Advertising to Gen-Z college students with memes? A focus group study. *Qualitative Market Research*, *27*(1), 1–18. Scopus. https://doi.org/10.1108/QMR-05-2023-0073

Vergeer, L., Soto, C., Bagnato, M., Pauzé, E., Amson, A., Ramsay, T., Olstad, D. L., Welch, V., & Potvin Kent, M. (2025). Examining differences in exposure to digital marketing of unhealthy foods reported by Canadian children and adolescents in two policy environments. *BMC Nutrition*, *11*(1). Scopus. https://doi.org/10.1186/s40795-025-01019-z

Whitton, C., Wong, Y. H. M., Lau, J., Chua, X. H., Müller, A. M., Tan, C. S., van Dam, R. M., Müller-Riemenschneider, F., & Rebello, S. A. (2024). Ecological momentary assessment of digital food and beverage marketing exposure and impact in young adults: A feasibility study. *Appetite*, *197*. Scopus. https://doi.org/10.1016/j.appet.2024.107338

Wisnujati, A., Wardana, L. K., Yudhanto, F., Anugrah, R. A., Fakhira, S., & Ogah, O. A. (2024). Application Technology and Digital Marketing for The Improvement of Production Capacity in Tempeh Small Industry. Dalam Rozaki Z., Alvarez Otero S., Ridzuan A.R., Mohd Kanafiah S.N.M., Zajadacz A., Triyono null, & Nurkomar I. (Ed.), *BIO. Web. Conf.* (Vol. 144). EDP Sciences; Scopus. https://doi.org/10.1051/bioconf/202414405001

You, S., Christofi, K., Tsappi, E., & Papageorgiou, G. (2024). Developing a Framework for City Brand-Image Promotion via Social Media Communication. Dalam Arami M., Baudier P., & Chang V. (Ed.), *Proc. Int. Conf. Financ., Econ., Manag. IT Bus., FEMIB* (hlm. 60–68). SciTePress; Scopus. https://doi.org/10.5220/0012699400003717

First Publication Right :

© Journal International of Lingua and Technology

This article is under: