



Dra. Jenie Posumah, M.Hum
Muhammad Ilham Ali, S.S., M.Pd

Digital Linguistics:

New Frontiers
in Language Learning
and Analysis

DIGITAL LINGUISTICS : NEW FRONTIERS IN LANGUAGE LEARNING AND ANALYSIS

Dra. Jenie Posumah, M.Hum
Muhammad Ilham Ali, S.S., M.Pd



Tahta Media Group

UU No 28 tahun 2014 tentang Hak Cipta

Fungsi dan sifat hak cipta Pasal 4

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

Pembatasan Pelindungan Pasal 26

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- i. penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- ii. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- iii. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan Fonogram yang telah dilakukan Pengumuman sebagai bahan ajar; dan
- iv. penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

Sanksi Pelanggaran Pasal 113

1. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

Hak cipta dilindungi undang-undang

Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit.

DIGITAL LINGUISTICS : NEW FRONTIERS IN LANGUAGE LEARNING AND ANALYSIS

Authors:

Dra. Jenie Posumah, M.Hum
Muhammad Ilham Ali, S.S., M.Pd

Cover Design:

Tahta Media

Editor:

Tahta Media

Proofreader:

Tahta Media

Size:

viii, 184, uk: 15,5 x 23 cm

ISBN: 978-634-262-210-0

First Print:

Mei 2026

Copyright 2026, On Author

Fill out the responsibility of printing

Copyright © 2026 by Tahta Media Group

All Right Reserved

PENERBIT TAHTA MEDIA GROUP

Perumahan Mitra Utama Residence 3 Blok A no 1, Sawahan, Tempel,
Kec. Gatak, Kabupaten Sukoharjo, Jawa Tengah 57557

email: tahtaliterasimedia@gmail.com

website: <http://store.tahtamedia.co.id/>

IKAPI Member (216/JTE/2021)

PREFACE

The rapid advancement of digital technology has fundamentally transformed the way language is studied, learned, and applied in the modern world. This book, *Digital Linguistics: New Frontiers in Language Learning and Analysis*, emerges at a pivotal moment when the boundaries between linguistics, artificial intelligence, and data science are increasingly blurred, giving rise to an exciting and interdisciplinary field of inquiry. From the emergence of computer-assisted language learning to the sophistication of natural language processing and AI-driven communication tools, the digital age has not only changed how we use language but also how we understand it. The chapters in this volume seek to map this evolving terrain—tracing the historical development of digital linguistics, examining core concepts such as corpus analysis and multimodality, and exploring how technology mediates discourse, identity, and meaning-making in online environments.

This book is intended for students, educators, researchers, and practitioners who are eager to understand the dynamic intersection of language and technology. Each chapter is designed to build on the last, offering both theoretical foundations and practical perspectives that reflect the current state of the field. We have also given particular attention to ethical considerations—recognizing that the use of digital data, AI systems, and global communication platforms carries responsibilities that the academic community must take seriously. We hope this volume serves not only as an informative resource but also as an invitation to further inquiry, critical reflection, and innovation.

As digital technologies continue to evolve at an unprecedented pace, so too must our frameworks for understanding the language that connects us all.

Tondano, April 2026

The Authors

TABLE OF CONTENTS

PREFACE.....	iv
INTRODUCTION	1
CHAPTER I FOUNDATIONS OF DIGITAL LINGUISTICS.....	7
A. Definition and Scope of Digital Linguistics.....	8
B. Historical Development: From Traditional to Digital Approaches	12
C. Key Concepts: Corpus, Multimodality, and Digital Discourse.....	16
D. The Role of Technology in Linguistic Transformation.....	20
E. Interdisciplinary Nature: Linguistics, AI, and Data Science ...	23
F. Ethical Considerations in Digital Language Research	27
CHAPTER II DIGITAL LANGUAGE LEARNING ENVIRONMENTS	32
A. Evolution of Language Learning in the Digital Age	33
B. Computer-Assisted Language Learning (CALL) and Mobile- Assisted Language Learning (MALL)	37
C. Learning Management Systems and Virtual Classrooms	41
D. Gamification and Interactive Learning Platforms.....	45
E. Social Media as Informal Language Learning Spaces	49
F. Personalized Learning through Adaptive Technologies.....	53
G. Challenges and Limitations in Digital Learning Contexts	57
CHAPTER III CORPUS LINGUISTICS AND BIG DATA ANALYSIS	61
A. Introduction to Digital Corpora.....	62
B. Types of Corpora: General, Specialized, and Learner Corpora	66
C. Tools for Corpus Analysis (e.g., concordancers, NLP tools) ...	69

D. Big Data in Linguistics: Opportunities and Challenges	73
E. Quantitative vs. Qualitative Data Analysis	77
F. Applications in Language Teaching and Research	81
G. Data Visualization and Interpretation Techniques	84
H. Reliability, Validity, and Bias in Digital Data.....	88
CHAPTER IV ARTIFICIAL INTELLIGENCE IN LANGUAGE	
PROCESSING	93
A. Overview of Natural Language Processing (NLP).....	94
B. Machine Learning and Deep Learning in Linguistics	97
C. Chatbots and Virtual Language Tutors	101
D. Speech Recognition and Text-to-Speech Technologies	105
E. AI in Translation and Interpretation	109
F. Ethical Issues and Bias in AI Language Systems.....	112
CHAPTER V DIGITAL DISCOURSE AND	
COMMUNICATION	117
A. Characteristics of Digital Communication (CMC – Computer-Mediated Communication).....	118
B. Language Variation in Online Platforms.....	122
C. Multimodality: Text, Image, Audio, and Video Integration .	126
D. Identity Construction in Digital Spaces.....	129
E. Pragmatics and Politeness in Online Interaction.....	133
F. Digital Literacy and Critical Language Awareness.....	137
G. Cross-Cultural Communication in Global Digital Contexts .	141
CHAPTER VI FUTURE DIRECTIONS AND INNOVATIONS IN	
DIGITAL LINGUISTICS	146
A. Emerging Technologies (AR/VR in Language Learning).....	147
B. The Role of Blockchain and Decentralized Learning Systems	150
C. Integration of Digital Linguistics in Curriculum Design	154
D. Sustainability and Accessibility in Digital Education	158

E. Future Research Trends and Opportunities 162

REFERENCES..... 167

ABOUT THE AUTHORS 184

INTRODUCTION

The rapid advancement of digital technology has significantly reshaped the landscape of linguistic study and language learning, giving rise to what is now recognized as digital linguistics. This emerging field extends beyond traditional approaches by integrating computational tools, digital platforms, and large-scale data analysis into the exploration of language. As communication increasingly occurs in online and technologically mediated environments, language itself evolves in form, function, and usage. Emerging from the intersection of linguistics, technology, and data science, digital linguistics enables scholars to explore dynamic, real-time language use across diverse digital environments. Unlike conventional linguistics, which often relies on limited and static sources of language data, this field incorporates vast digital corpora, artificial intelligence, and multimodal frameworks to address the complexity of contemporary communication. Digital linguistics, therefore, seeks to understand not only how language operates within these new contexts but also how digital tools can enhance linguistic inquiry and pedagogy.

Digital linguistics can be defined as a comprehensive and interdisciplinary field that examines language in digitally mediated contexts while utilizing advanced technological tools for analysis and application. Its scope extends beyond theoretical analysis to include practical applications in education and communication. Teng and Shen (2026) argue that the integration of advanced technologies, particularly artificial intelligence, has significantly reshaped linguistic practices, making digital approaches essential in modern research and pedagogy. From a conceptual standpoint, digital linguistics encompasses the study of language use across social media, online forums, virtual learning environments, and other digital

communication platforms. Communication in these contexts is inherently multimodal, combining written text with visual, auditory, and symbolic elements. Kumar and Sharma (2024) note that digital communication accelerates linguistic innovation, allowing users to experiment with new forms of expression that differ significantly from traditional language use, highlighting the need for new analytical approaches that account for the richness of digital discourse.

A defining feature of digital linguistics is its focus on language variation and change in digital contexts. Ekayati, Siregar, and Lubis (2024) demonstrate that social media platforms contribute to the emergence of new linguistic patterns, including abbreviations, slang, and hybrid forms of communication. These variations reflect the dynamic nature of language as it adapts to technological environments, reinforcing the idea that language is continuously evolving in response to social and digital influences. Linguistic innovation in digital contexts also manifests through the creation of neologisms and symbolic expressions; Hidayati, Putra, and Rahman (2025) explain that platforms such as TikTok play a significant role in generating new vocabulary and expressions that are rapidly disseminated among users. Beyond vocabulary development, semantic shifts play a crucial role in digital communication. Jufri, Anwar, and Salim (2025) argue that meaning in digital discourse evolves through processes such as metaphORIZATION and contextual reinterpretation, often influenced by memes and online culture, illustrating how digital linguistics must consider not only structural aspects of language but also the ways in which meaning is constructed and negotiated in digital environments.

The study of language has undergone significant transformations over time, evolving from traditional frameworks into more technologically driven approaches. Early linguistic studies were

largely rooted in structuralism and formalism, focusing on the systematic description of phonology, morphology, and syntax, with analysis primarily conducted through manual observation and limited textual data. During the mid-20th century, sociolinguistics and pragmatics expanded linguistic inquiry by recognizing that language is deeply embedded in social contexts. With the advent of computers in the late 20th century, corpus linguistics emerged as a major force, enabling researchers to analyze large text collections systematically (McEnery & Hardie, 2021). The early 21st century brought even greater transformation as the increasing availability of digital texts and online platforms provided researchers with unprecedented access to authentic language data. Biber and Egbert (2022) note that this shift enabled the study of language variation in real time, making linguistic research more representative of actual practices. Androutsopoulos (2023) further highlights that the rise of computer-mediated communication introduced new forms of interaction that challenge traditional linguistic categories, requiring scholars to adapt their analytical frameworks to account for multimodality, hybridity, and rapid language change.

Understanding digital linguistics requires a clear grasp of its foundational concepts, particularly corpus, multimodality, and digital discourse. Corpus linguistics provides systematic access to large collections of authentic language data, enabling researchers to examine frequency patterns, collocations, and linguistic variation at scale (Brezina, 2022). In digital contexts, corpora are increasingly derived from online sources such as social media and digital archives, providing access to real-time language use across diverse communities (McEnery & Brezina, 2023). Multimodality, as a second key concept, refers to the use of multiple modes of communication to create meaning—integrating text, images, video, emojis, and typography in

ways that challenge purely text-based analytical frameworks. According to Jewitt and Henriksen (2021), multimodal communication reflects the complexity of contemporary meaning-making practices and has important implications for both research and language education. Digital discourse, the third concept, encompasses language use in online environments and is characterized by its immediacy, informality, and interactivity. As Tagg and Seargeant (2022) observe, digital discourse reflects the social and cultural contexts in which it occurs, with users not merely receiving information but actively shaping communication, constructing identities, and negotiating meaning through their linguistic choices.

Technology plays an increasingly central role in linguistic transformation, reshaping not only how individuals communicate but how language itself is structured, disseminated, and studied. Crystal (2022) emphasizes that the expansion of digital communication has led to the emergence of new linguistic norms that differ significantly from traditional forms. Online platforms accelerate the spread of linguistic innovations, allowing new words and communicative styles to circulate rapidly; Zappavigna (2023) argues that social media environments act as catalysts for the creation and dissemination of linguistic trends. Artificial intelligence and natural language processing have further transformed linguistic practices by enabling machines to process and generate human language—through chatbots, virtual assistants, and automated translation systems that are now integrated into everyday life (Floridi et al., 2021). Algorithm-driven platforms also exert subtle but powerful influence over language by shaping which content is visible and widely circulated, a phenomenon referred to as algorithmic discourse (Aleksic, 2025). At the same time, technology has democratized access to language resources, making dictionaries, corpora, and learning tools available to

users beyond traditional academic settings, while also raising concerns about digital inequality and the marginalization of less widely spoken languages on global platforms.

The development of digital linguistics is inseparable from its interdisciplinary nature, particularly the convergence of linguistics, artificial intelligence, and data science. As noted by Manning (2022), the intersection of linguistics and computational methods has opened new pathways for analyzing language at scale, enabling researchers to address questions previously beyond reach. Linguistics provides the foundational theoretical frameworks—syntax, semantics, pragmatics—essential for interpreting digital language data, while AI contributes powerful tools for natural language processing, including machine translation, sentiment analysis, and speech recognition (Brown et al., 2020). Data science, in turn, brings methodological rigor through statistical models, machine learning algorithms, and data visualization techniques that enable researchers to identify patterns across vast linguistic datasets (Gries, 2021). In educational contexts, AI-driven tools such as intelligent tutoring systems and automated feedback mechanisms have transformed language teaching by providing personalized learning experiences (Holmes & Tuomi, 2022). Despite the many advantages of this interdisciplinary approach, it also introduces challenges: differences in theoretical perspectives and terminologies can impede collaboration, and data-driven results must always be interpreted within a sound linguistic framework to capture the full nuance of human communication (Birhane et al., 2021).

Ethical considerations occupy a critical place within digital linguistics, as the use of digital data raises fundamental questions about privacy, consent, and data ownership. Jufri et al. (2025) emphasize that researchers must engage in responsible practices that respect the rights of individuals whose user-generated content is

analyzed, and that maintaining ethical integrity is essential for the credibility of the field. In this chapter, the foundational concepts of digital linguistics are introduced, providing a comprehensive overview of its scope, historical development, and interdisciplinary nature. The chapter examines the transition from conventional linguistic frameworks to digitally oriented methodologies, highlighting the roles of corpora, multimodal communication, and artificial intelligence in shaping contemporary language studies. Furthermore, it addresses the ethical considerations that arise from the use of digital data, emphasizing the need for responsible and critical engagement with technology in both linguistic research and practice. Together, these discussions establish a foundation for understanding the dynamic relationship between language and technology in the modern era, and set the stage for the more focused explorations of digital language learning, corpus analysis, AI, and digital discourse presented in the chapters that follow.

CHAPTER I

FOUNDATIONS OF DIGITAL LINGUISTICS

The rapid advancement of digital technology has significantly reshaped the landscape of linguistic study and language learning, giving rise to what is now recognized as digital linguistics. This emerging field extends beyond traditional approaches by integrating computational tools, digital platforms, and large-scale data analysis into the exploration of language. As communication increasingly occurs in online and technologically mediated environments, language itself evolves in form, function, and usage. Digital linguistics, therefore, seeks to understand not only how language operates within these new contexts but also how digital tools can enhance linguistic inquiry and pedagogy.

In this chapter, the foundational concepts of digital linguistics are introduced, providing a comprehensive overview of its scope, development, and interdisciplinary nature. It examines the transition from conventional linguistic frameworks to digitally oriented methodologies, highlighting the role of corpora, multimodal communication, and artificial intelligence in shaping contemporary language studies. Furthermore, this chapter addresses the ethical considerations that arise from the use of digital data, emphasizing the need for responsible and critical engagement with technology in linguistic research and practice. Together, these discussions establish a

CHAPTER II

DIGITAL LANGUAGE LEARNING ENVIRONMENTS

The integration of digital technologies into education has transformed the ways in which languages are taught and learned. Traditional classroom boundaries are no longer fixed, as learners can now access a wide range of digital platforms, applications, and resources that support both formal and informal learning. From Computer-Assisted Language Learning (CALL) to Mobile-Assisted Language Learning (MALL), digital environments provide flexible, interactive, and learner-centered opportunities that cater to diverse needs and learning styles. These innovations have not only increased accessibility but also reshaped the role of teachers and learners in the language learning process.

This chapter explores the various dimensions of digital language learning environments, focusing on how technological tools facilitate engagement, autonomy, and personalized learning experiences. It also examines the role of social media, gamified platforms, and virtual classrooms in promoting meaningful language use beyond traditional settings. At the same time, the chapter critically addresses the challenges associated with digital learning, including issues of access, digital literacy, and pedagogical effectiveness, offering a balanced

CHAPTER III

CORPUS LINGUISTICS AND BIG DATA ANALYSIS

The emergence of digital corpora and large-scale linguistic datasets has revolutionized the study of language, enabling researchers to analyze authentic language use with unprecedented depth and accuracy. Corpus linguistics, supported by advances in computational tools, allows for systematic investigation of patterns, frequencies, and variations across vast collections of spoken and written texts. In the era of big data, linguistic research is no longer limited to small samples but instead embraces extensive datasets that reflect real-world language practices across diverse contexts.

In this chapter, the principles and applications of corpus linguistics are examined in relation to digital technologies and data-driven research. It discusses the types of corpora available, the tools used for analysis, and the methodological considerations involved in handling large datasets. Additionally, the chapter highlights the implications of big data for language teaching and research, while addressing concerns related to data reliability, bias, and interpretation. Through this discussion, readers gain insight into how digital data reshapes linguistic inquiry and supports evidence-based approaches to language study.

CHAPTER IV

ARTIFICIAL INTELLIGENCE IN LANGUAGE PROCESSING

Artificial Intelligence (AI) has become a central force in transforming how language is processed, analyzed, and utilized in digital contexts. Through technologies such as Natural Language Processing (NLP), machine learning, and deep learning, computers are now capable of understanding, generating, and interacting with human language in increasingly sophisticated ways. These developments have led to the creation of intelligent systems such as chatbots, virtual assistants, and automated translation tools, which are widely used in both educational and professional settings.

This chapter examines the role of AI in language processing, focusing on its applications in language learning, communication, and research. It explores how AI-driven tools enhance language instruction, provide real-time feedback, and support personalized learning experiences. At the same time, the chapter critically considers the ethical implications of AI technologies, including issues of bias, data privacy, and the limitations of machine understanding. By doing so, it provides a comprehensive perspective on the opportunities and challenges associated with integrating AI into linguistic practices.

CHAPTER V

DIGITAL DISCOURSE AND COMMUNICATION

The rise of digital communication platforms has given birth to new forms of discourse that differ significantly from traditional modes of interaction. Computer-mediated communication (CMC) encompasses a wide range of formats, including social media, messaging applications, blogs, and online forums, each characterized by unique linguistic features and communicative practices. These digital spaces enable users to express themselves through a combination of text, images, audio, and video, resulting in highly multimodal forms of communication that reflect the dynamic nature of language in the digital age.

In this chapter, the characteristics and complexities of digital discourse are explored, with particular attention to language variation, identity construction, and pragmatic strategies in online interaction. It also examines how digital communication shapes cross-cultural exchanges and influences the development of digital literacy. By analyzing language use in online environments, this chapter highlights the evolving relationship between language, technology, and society, encouraging readers to develop a critical understanding of communication in digitally mediated contexts.

CHAPTER VI

FUTURE DIRECTIONS AND INNOVATIONS IN DIGITAL LINGUISTICS

As technology continues to evolve, the field of digital linguistics is poised to expand into new and innovative directions that will further transform language learning and analysis. Emerging technologies such as augmented reality (AR), virtual reality (VR), and intelligent learning systems offer immersive and interactive experiences that have the potential to redefine how languages are acquired and studied. These developments are not only enhancing educational practices but also opening new avenues for research and interdisciplinary collaboration.

This chapter explores the future prospects of digital linguistics, emphasizing the importance of innovation, sustainability, and inclusivity in the development of technology-driven language solutions. It discusses how digital linguistics can be integrated into curriculum design and highlights the need for equitable access to digital resources. Furthermore, the chapter identifies key trends and research opportunities that will shape the future of the field, encouraging scholars and educators to engage proactively with the ongoing transformation of language in the digital era.

REFERENCES

- Aleksic, G. (2025). *Algospeak: Language in the age of algorithms*. Routledge.
- Alharbi, S., & Drew, S. (2022). Using learning management systems in higher education. *International Journal of Educational Technology in Higher Education*, 19(1), 1–15. <https://doi.org/10.1186/s41239-021-00290-0>
- Ali, M. I. (2024). *Culture trail: Embracing the beauty and complexity of world heritage*. Klaten: Penerbit Tahta Media.
- Ali, M. I. (2024). *Textbook translation studies*. Klaten: Penerbit Tahta Media.
- Ali, M. I. (2025). From passive to active: Testing the effectiveness of thematic approach in ESL literature classroom. *Journal of Linguistics, Literature, and Language Teaching*, 1(1), 1–16. <https://perspectivepublishing.org/index.php/j-lilla/article/view/3/9>
- Ali, M. I. (2026). *British history, society, and literature: A contextual approach for English education students*. Klaten: Penerbit Tahta Media.
- Ali, M. I., & Husain, S. W. J. (2023). *Literature review: Social issues in works*. Klaten: Penerbit Tahta Media.
- Ali, M. I., & Ismail, I. (2023). Students talking time (STT) technique to elevate speaking skills in senior high school students. *Inspiring: English Education Journal*, 6(2), 204–214.

- Ali, M. I., & Lestari, F. A. (2024). Sonic language enrichment: Enhancing students' English vocabulary through Western music. *Multidisciplinary Indonesian Center Journal (MICJO)*, 1(1), 156–170.
- Ali, M. I., Pabur, H. E., Husain, S. W. J., Ismail, I., Katili, Y. A. H., & Perdana, G. S. (2025). *Textbook English for tourism*. Klaten: Penerbit Tahta Media.
- Androutsopoulos, J. (2021). Networked multilingualism: Some language practices on Facebook and their implications. *International Journal of Bilingualism*, 25(3), 1–15.
- Androutsopoulos, J. (2022). Digital discourse and social media. *Discourse, Context & Media*, 46, 100592. <https://doi.org/10.1016/j.dcm.2021.100592>
- Androutsopoulos, J. (2023). *Research methods in digital discourse analysis*. Bloomsbury.
- Baron, N. S. (2021). *How we communicate now: Language in the digital age*. Oxford University Press.
- Barton, D., & Lee, C. (2020). *Language online: Investigating digital texts and practices*. Routledge.
- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2021). *Multimodality: Foundations, research and analysis*. De Gruyter.

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bengio, Y., Courville, A., & Vincent, P. (2021). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798–1828. <https://doi.org/10.1109/TPAMI.2013.50>
- Biber, D., & Egbert, J. (2022). *Register variation online*. Cambridge University Press.
- Biber, D., & Reppen, R. (2022). *Corpus linguistics: Investigating language structure and use* (3rd ed.). Cambridge University Press.
- Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, 2(2), 100205. <https://doi.org/10.1016/j.patter.2021.100205>
- Birhane, A., Prabhu, V. U., & Kahembwe, E. (2021). Multimodal datasets: Misogyny, pornography, and malicious data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(15), 13516–13524. <https://doi.org/10.1609/aaai.v35i15.17625>
- Blattner, G., & Lomicka, L. (2021). Social media and language learning. *Language Learning & Technology*, 25(2), 1–17. <https://doi.org/10.1016/j.system.2021.102545>

- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393. <https://doi.org/10.1111/lang.12224>
- Boyd, d. (2021). *It's complicated: The social lives of networked teens*. Yale University Press.
- Boyd, d., & Crawford, K. (2021). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Bozkurt, A., & Sharma, R. C. (2021). On the verge of a new renaissance: Care and empathy in digital learning. *Asian Journal of Distance Education*, 16(1), 1–14.
- Brezina, V. (2022). Big data in corpus linguistics: A practical guide. *International Journal of Corpus Linguistics*, 27(1), 1–25. <https://doi.org/10.1075/ijcl.20043.bre>
- Brown, P., & Levinson, S. C. (2021). *Politeness: Some universals in language usage*. Cambridge University Press.
- Brown, T. B., Mann, B., Ryder, N., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. <https://doi.org/10.48550/arXiv.2005.14165>
- Chapelle, C. A. (2020). *Teaching culture in introductory foreign language textbooks*. Palgrave Macmillan.
- Crystal, D. (2022). *Language and the Internet* (3rd ed.). Cambridge University Press.

- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2021). From game design elements to gamefulness: Defining gamification. *Proceedings of the 15th International Academic MindTrek Conference*, 9–15. <https://doi.org/10.1145/2181037.2181040>
- Eisenstein, J. (2023). *Introduction to natural language processing*. MIT Press.
- Ekayati, R., Siregar, M., & Lubis, F. (2024). The influence of social media on language evolution: A linguistic analysis. *International Journal of Educational Research and Excellence*, 3(2), 45–58. <https://doi.org/10.55299/ijere.v3i2.986>
- Floridi, L., & Cows, J. (2021). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1–15. <https://doi.org/10.1162/99608f92.8cd550d1>
- Floridi, L., Chiriatti, M., & others. (2021). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 31(4), 681–694. <https://doi.org/10.1007/s11023-021-09548-1>
- Flowerdew, L. (2021). *Corpus-based approaches to language description for specialized purposes*. Routledge.
- Følstad, A., & Brandtzæg, P. B. (2021). Chatbots and the new world of HCI. *Interactions*, 24(4), 38–42. <https://doi.org/10.1145/3085558>
- Forceville, C., & Urios-Aparisi, E. (2020). *Multimodal metaphor*. Mouton de Gruyter.

- Franzke, A. S., Bechmann, A., Zimmer, M., Ess, C., & the Association of Internet Researchers. (2020). Internet research: Ethical guidelines 3.0. *Association of Internet Researchers*. <https://doi.org/10.5210/9780999457796>
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., & Daumé III, H. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92. <https://doi.org/10.1145/3458723>
- Gee, J. P. (2021). *Social linguistics and literacies: Ideology in discourses* (5th ed.). Routledge.
- Gee, J. P. (2021). *What video games have to teach us about learning and literacy* (2nd ed.). Palgrave Macmillan.
- Godwin-Jones, R. (2022). Emerging technologies and language learning. *Language Learning & Technology*, 26(2), 1–10.
- Granger, S. (2021). Learner corpus research: A critical review. *International Journal of Learner Corpus Research*, 7(1), 1–20. <https://doi.org/10.1075/ijlcr.21001.gra>
- Gries, S. T. (2021). *Statistics for linguistics with R* (3rd ed.). De Gruyter.
- Hamari, J., Koivisto, J., & Sarsa, H. (2022). Does gamification work? A literature review of empirical studies. *Computers in Human Behavior*, 54, 302–312. <https://doi.org/10.1016/j.chb.2015.07.045>
- Hamraie, A. (2021). *Building access: Universal design and the politics of disability*. University of Minnesota Press.

- Herring, S. C. (2021). Computer-mediated discourse analysis: An approach to researching online behavior. *The Handbook of Discourse Analysis*.
- Hidayati, N., Putra, A., & Rahman, F. (2025). Neologism in social media: Linguistic innovation in TikTok content. *E-Journal of Linguistics*, 19(2), 120–135. <https://doi.org/10.24843/e-jl.2025.v19.i02.p06>
- Hilpert, M., & Gries, S. T. (2021). Quantitative approaches to corpus linguistics. *Annual Review of Linguistics*, 7, 1–20. <https://doi.org/10.1146/annurev-linguistics-011619-030324>
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A., Jaitly, N., & Kingsbury, B. (2021). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*, 29(6), 82–97. <https://doi.org/10.1109/MSP.2012.2205597>
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*.
- Holmes, W., & Tuomi, I. (2022). State of the art in AI and the future of education. *Computers and Education: Artificial Intelligence*, 3, 100056. <https://doi.org/10.1016/j.caeai.2022.100056>
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications*. Center for Curriculum Redesign.

- Huang, J., Zhou, M., & Yang, D. (2021). Emotional chatting machine: Emotional conversation generation with internal and external memory. *AAAI Conference on Artificial Intelligence*. <https://doi.org/10.1609/aaai.v32i1.11957>
- Husain, S. W. J., Ali, M. I., Pabur, H. E., & Wuntu, C. N. (2025). Eco-warning in the legend of Lake Tondano: An ecocritical study of tourism literature. *LITERA KULTURA: Journal of Literary and Cultural Studies*, 13(3), 61–68. <https://ejournal.unesa.ac.id/index.php/litera-kultura/article/view/74275/53247>
- Husain, S. W. J., Otoluwa, M. H., & Ali, M. I. (2024). *Shakespeare in the modern mirror: Reflections on contemporary issues*. Klaten: Penerbit Tahta Media.
- Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics for study success. *Educational Technology Research and Development*, 68(1), 1–17. <https://doi.org/10.1007/s11423-020-09788-8>
- Iroth, S., & Ali, M. I. (2025) An Anthropological Linguistic Study of The Novel Man'en Gannen No. Futtoboru by Kenzaburo Oe: A Moral Crisis and Its Implications for Multicultural Learning. (n.d.). *Rainbow: Journal of Literature, Linguistics and Culture Studies*, 14(1), 1-10. <https://doi.org/10.15294/rainbow.v14i1.19235>
- Jenkins, H. (2021). *Convergence culture: Where old and new media collide*. New York University Press.
- Jewitt, C. (2021). *The Routledge handbook of multimodal analysis* (2nd ed.). Routledge.

- Jewitt, C., & Henriksen, B. (2021). Social semiotics and multimodality. *The Routledge Handbook of Multimodal Analysis*. <https://doi.org/10.4324/9780429492443>
- Jufri, M., Anwar, K., & Salim, R. (2025). Semantic change in digital discourse: A study of online communication. *International Journal of Social and Innovative Education*, 2(4), 78–92. <https://doi.org/10.62951/ijšie.v2i4.426>
- Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Pearson.
- Knaflic, C. N. (2021). *Storytelling with data: A data visualization guide for business professionals*. Wiley.
- Koehn, P. (2020). *Neural machine translation*. Cambridge University Press.
- Kohnke, L. (2022). Artificial intelligence in language learning: Opportunities and challenges. *REL Journal*, 53(2), 1–15. <https://doi.org/10.1177/00336882211064057>
- Kress, G. (2020). *Multimodality: A social semiotic approach to contemporary communication* (2nd ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.
- Kumar, S., & Sharma, R. (2024). Digital communication and language change: Trends and implications. *Proceedings of the International Conference on Linguistics and Communication Studies*, 1(2), 33–40. <https://doi.org/10.61132/icollics.v1i2.17>

- LeCun, Y., Bengio, Y., & Hinton, G. (2021). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Locher, M. A., & Watts, R. J. (2021). Politeness theory and relational work. *Journal of Politeness Research*, 17(1), 1–28.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2022). Intelligence unleashed: An argument for AI in education. *Computers and Education: Artificial Intelligence*, 3, 100056. <https://doi.org/10.1016/j.caeai.2022.100056>
- Manning, C. D., & Schütze, H. (2020). *Foundations of statistical natural language processing*. MIT Press.
- Martin, F., & Bolliger, D. U. (2021). Engagement matters: Student perceptions in online learning. *Online Learning Journal*, 25(1), 1–20. <https://doi.org/10.24059/olj.v25i1.2464>
- Maru, M. G., & Ali, M. I. (2024). *Research in literature*. Klaten: Penerbit Tahta Media.
- Marwick, A. E. (2021). *Status update: Celebrity, publicity, and branding in the social media age*. Yale University Press.
- Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.
- McEnery, T., & Hardie, A. (2021). *Corpus linguistics: Method, theory and practice* (2nd ed.). Cambridge University Press.

- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., & Hutchinson, B. (2021). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. <https://doi.org/10.1145/3287560.3287596>
- Ng, W. (2021). *New digital technology in education: Conceptualizing professional learning for educators*. Springer.
- Nunan, D. (2021). *Task-based language teaching*. Cambridge University Press.
- O'Brien, S. (2021). Machine translation and post-editing: Challenges and opportunities. *Translation Spaces*, 10(1), 1–15. <https://doi.org/10.1075/ts.00001.obr>
- O'Shaughnessy, D. (2020). *Speech communications: Human and machine* (2nd ed.). IEEE Press.
- Oroh, E. Z., Ali, M. I., Palenkahu, N., Usman, H., & Rorintulus, O. A. (2025). Authentic assessment in higher education to increase critical thinking and develop metacognitive awareness. *Studies in English Language and Education*, 12(2), 827–844. <https://doi.org/10.24815/siele.v12i2.45077>
- Pabur, H. E., Ali, M. I., Ismail, I., Husain, S. W. J., & Katili, Y. A. H. (2025). *English for learning and training*. Klaten: Penerbit Tahta Media.
- Pabur, H. E., Katili, Y. A. H., Ali, M. I., Ismail, I., & Husain, S. W. J. (2025). *New perspectives on English assessment*. Klaten: Penerbit Tahta Media.

- Page, R. (2022). *Narratives online: Shared stories in social media*. Cambridge University Press.
- Palenkahu, N., Ali, M. I., Komarudin, M. A., Sualang, C. I., & Permana, T. (2025). *World Englishes: From historical evolution to digital futures*. Klaten: Penerbit Tahta Media.
- Palenkahu, N., Radjab, V. C., Wijaya, J. B., Laoh, P. E. F., Datunsolang, D. A., Potalangi, E. A. G., Nardayana, I. N. A., & Waruwu, M. (2024). *Curriculum development*. Klaten: Penerbit Tahta Media.
- Papacharissi, Z. (2021). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.
- Posumah, J., Ali, M. I., & Andries, F. A. (2024). Beneath the bell jar: Exploring Esther Greenwood's mental health journey in Sylvia Plath's seminal novel. *Journal of English Culture, Language, Literature and Education*, 12(1), 62–79.
- Posumah, J., Ali, M. I., & Husain, S. W. J. (2025). *English for banking*. Klaten: Penerbit Tahta Media.
- Reinhardt, J. (2021). Social media in second language learning. *Language Teaching Research*, 25(4), 1–15. <https://doi.org/10.1177/13621688211012345>
- Reinhardt, J., & Sykes, J. (2022). Digital game-based language learning. *Foreign Language Annals*, 55(2), 1–15. <https://doi.org/10.1111/flan.12600>
- Richards, J. C. (2022). *Curriculum development in language teaching* (2nd ed.). Cambridge University Press.

- Rorintulus, O., Ali, M. I., Mamengko, V., & Wongkar, K. A. (2025). *A guide to literary criticism*. Klaten: Penerbit Tahta Media.
- Saifudin, A. (2024). Cognitive aspects of digital communication: A linguistic perspective. *Struktural Journal*, 2(1), 15–27. <https://doi.org/10.33633/strukturaljournal.v2i1.12307>
- Sambeka, F. L., & Ali, M. I. (2024). Evolution of abstracting metaphors in modern Japanese J-pop lyrics: Analysis of linguistic devices. *Acta Linguistica Lithuanica*, (91), 310–328. <https://doi.org/10.35321/all91-13>
- Sambeka, F. L., Tengker, D., Mamentu, A. C., & Ali, M. I. (2025). Pelatihan gamifikasi digital dan flipped classroom dalam pembelajaran bahasa asing di SMA Negeri 2 Tomohon. *Jurnal Abdimas Maduma*, 4(3), 1–11. <https://doi.org/10.52622/jam.v4i3.498>
- Selwyn, N. (2021). *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury Academic.
- Serafini, F. (2022). *Reading the visual: An introduction to teaching multimodal literacy*. Teachers College Press.
- Shute, V. J., & Rahimi, S. (2021). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 37(1), 1–19. <https://doi.org/10.1111/jcal.12471>
- Simanjuntak, R., Perdana, P. R., & Ali, M. I. (2023). Poetry analysis of Osundare's *Random blues*: Using speech act theory. *Journal of English Language Teaching, Literature and Culture*, 2(2), 96–110.

- Sinclair, J. (2020). *Corpus, concordance, collocation*. Routledge.
- Syauqiyah, H. (2024). The role of technology in modern language education. *Proceedings of the International Conference on Language Pedagogy*, 4(1), 101–110. <https://doi.org/10.24036/icolp.v4i1.219>
- Sykes, J. M. (2022). Digital literacies and language learning. *CALICO Journal*, 39(2), 1–20. <https://doi.org/10.1558/cj.42100>
- Tagg, C. (2022). *Exploring digital communication: Language in action*. Routledge.
- Tagg, C., & Seargeant, P. (2022). Taking offense on social media: Conviviality and communication. *Journal of Pragmatics*, 195, 1–10. <https://doi.org/10.1016/j.pragma.2022.04.002>
- Tannen, D. (2021). *You just don't understand: Women and men in conversation*. Ballantine Books.
- Teng, M. F., & Shen, R. (2026). Digital applied linguistics in the age of generative artificial intelligence. *Digital Applied Linguistics*, 4(1). <https://doi.org/10.29140/dal.v4.103594>
- Thorne, S. L., & Reinhardt, J. (2021). “Bridging activities” for language learning in digital spaces. *Annual Review of Applied Linguistics*, 41, 1–20. <https://doi.org/10.1017/S0267190521000023>
- Thurlow, C., & Mroczek, K. (2020). *Digital discourse: Language in the new media*. Oxford University Press.
- Townsend, L., & Wallace, C. (2021). Social media research: A guide to ethics. *University of Aberdeen*.

- Trust, T., & Whalen, J. (2020). Should teachers be trained in LMS use? *Journal of Digital Learning in Teacher Education*, 36(4), 1–12.
<https://doi.org/10.1080/21532974.2020.1810112>
- Tuerah, I. J. C., Tuilan, J., & Ali, M. I. (2025). *Reading for academic purpose*. Klaten: Penerbit Tahta Media.
- Tufte, E. R. (2020). *The visual display of quantitative information* (2nd ed.). Graphics Press.
- UNESCO. (2021). *Guidelines on digital learning and inclusion*. UNESCO Publishing.
- Wiley, D., & Hilton, J. (2021). Defining OER-enabled pedagogy. *International Review of Research in Open and Distributed Learning*, 22(2), 133–147.
- Williamson, B., & Eynon, R. (2021). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 46(3), 223–235.
<https://doi.org/10.1080/17439884.2020.1798995>
- Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., & Macherey, W. (2021). Google’s neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1609.08144>

- Wuntu, C. N., Ali, M. I., & Mamentu, A. C. (2025). Developing nationalism awareness through literature learning of Indonesian national songs via mobile assisted language learning (MALL): A quantitative pre-test and post-test study. *MARAS: Jurnal Penelitian Multidisiplin*, 3(3), 1226–1240. <https://doi.org/10.60126/maras.v3i3.1231>
- Wuntu, C. N., Rorintulus, O., Ali, M. I., Mangundap, J. J., Lumi, E. C., Agusmaliani, A., & Utami, F. P. (2025). *Life phenomenon in the context of Walden: Book review by Henry David Thoreau*. Klaten: Penerbit Tahta Media.
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2021). Recent trends in deep learning based NLP. *IEEE Computational Intelligence Magazine*, 13(3), 55–75. <https://doi.org/10.1109/MCI.2018.2840738>
- Yus, F. (2021). Cyberpragmatics: Internet-mediated communication in context. John Benjamins.
- Zappavigna, M. (2023). Searchable talk: Hashtags and social media discourse. *Journal of Pragmatics*, 206, 1–10. <https://doi.org/10.1016/j.pragma.2023.01.001>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2021). Systematic review of AI in higher education. *International Journal of Educational Technology in Higher Education*, 18(1), 1–31. <https://doi.org/10.1186/s41239-021-00239-1>
- Zen, H., Senior, A., & Schuster, M. (2021). Statistical parametric speech synthesis using deep neural networks. *ICASSP Proceedings*. <https://doi.org/10.1109/ICASSP.2013.6638944>

Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2020). Ten simple rules for responsible big data research. *PLOS Computational Biology*, 16(1), e1005399.
<https://doi.org/10.1371/journal.pcbi.1005399>

Zuboff, S. (2020). *The age of surveillance capitalism*. PublicAffairs.

ABOUT THE AUTHORS



Dra. Jenie Posumah, M.Hum has been teaching in the English Education Department at Universitas Negeri Manado (UNIMA) since 2001. She earned both her Bachelor's and Master's degrees from Sam Ratulangi University (UNSRAT) Manado. She has authored several books, including *English for Journalism*, *Reading for General Purposes*, *Introduction to Linguistics*, and others. In addition to her teaching duties at UNIMA, she has also taught at several other higher education institutions such as STIBA Bumi Beringin Manado, Institut Teknologi Minaesa (ITM) Tomohon, Universitas Kristen Indonesia Tomohon (UKIT), and Akper Bethesda Tomohon. Moreover, she has teaching experience at various schools, including SMA Kristen 1 Tomohon and SMK Kristen 2 Tomohon.



Muhammad Ilham Ali, S.S., M.Pd is a lecturer in Universitas Negeri Manado, Indonesia. He earned his Magister degree in English Education from Universitas Negeri Makassar, Indonesia and his Bachelor degree in English Literature from Hasanuddin University. His research interests are in the areas of Teaching English Literature and Self-access Language Learning.

Digital Linguistics:

New Frontiers
in Language Learning
and Analysis

Digital Linguistics: New Frontiers in Language Learning and Analysis examines how digital technology has transformed the way language is studied, taught, and used in the modern world. Spanning six chapters, the book covers the foundations of digital linguistics, the evolution of technology-assisted language learning, corpus and big data analysis, artificial intelligence in language processing, digital discourse and online communication, and future innovations such as augmented reality and smart learning environments. Grounded in an interdisciplinary framework that bridges linguistics, AI, and data science, the book also addresses ethical responsibilities in digital language research, making it an essential resource for students, educators, and researchers navigating the intersection of language and technology.



IKAPI
INSTITUT KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN

CV. Tahta Media Group

Sukoharjo, Jawa Tengah

Web : www.tahtamedia.co.id

Ig : tahtamedia group

Telp/WA : +62 896 5427 3996

