

CHAPTER I

INTRODUCTION

1.1. Background of the research

Corpus Linguistic is one of the most common approaches to studying language, especially in the context of grammatical patterns in natural language used in everyday life. Unlike traditional linguistic analysis, which emphasizes the identification of language structures and explains how the smallest units of language can be combined into grammatical units, corpus analysis emphasizes the use of language in everyday contexts.¹ In this study, corpus-based research will be used in the context of grammar patterns that frequently appear in news texts and textbooks to determine the core language patterns of news texts and textbooks with the aim of supporting the learning process of students, especially in comprehensive reading skills.

Corpus Linguistic itself has the following characteristics: 1) it is empirical because it analyzes language patterns commonly used in everyday life, 2) it uses a large collection of natural texts, which in this study will be referred to as a “corpus” as the basis for analysis, and 3) it uses computer analysis in both automatic and interactive forms with quantitative and qualitative approaches.² Therefore, rather than making assumptions based on related theories, this approach relies on data in the form of a corpus that will be considered for research in the context of how

¹ Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

² Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

language is used in speaking and writing, and computer-based research itself is able to reduce ambiguity.

The most fundamental concept in corpus-based linguistic research is the association of grammatical patterns, the systematic way in which linguistic features are used with other linguistic and non-linguistic features. Through language pattern association, there are two general types of problem formulations that can be used, namely: investigating the use of language features (both lexical and grammatical) and investigating the characteristics of a text or various types of texts.³ This approach allows researchers to examine general patterns used in natural language rather than assessing grammar itself, with two main research objectives: assessing the extent to which a pattern is found in corpus data and analyzing the contextual factors that influence the diversity of language patterns themselves.

Grammar itself can be understood as a systematic description of how words are organized into larger units such as phrases and clauses in a language.⁴ Rather than focusing solely on rules, descriptions of grammar aim to explain how linguistic forms are arranged and combined in actual language use. This perspective explains grammar as a resource in the formation of meaning, providing a structural foundation through which written and unwritten texts are formed and interpreted.

From a corpus perspective, grammar is investigated in the form of grammatical patterns, explained as recurring structured configurations that emerge

³ Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

⁴ Murray J. Munro. *Introduction to English Grammatical Description*. John Benjamins Publishing Company, 2008.

from the use of language commonly found in real life. Grammar patterns can be identified by analyzing how often they appear and how they are distributed through phrase and clause constructions in large amounts of text.⁵ This study uses the grammatical framework embedded in Longman Grammar of Spoken and Written English, which divides grammatical patterns into several categories based on their form and function in written texts and observes how certain patterns are associated with specific registers.

From a pedagogical perspective, grammar is not merely a set of prescriptive rules but a dynamic system that supports language comprehension and production. Grammatical competence is foundational for advanced language skills, particularly in academic contexts where texts often contain complex syntactic structures and dense informational content.⁶ For that reason, University-level EFL students must master various grammatical constructions such as embedded clauses, passive sentences, and nominalization commonly found in academic and scientific texts. The ability to decode and interpret these structures is essential for extracting meaning and engaging critically with scholarly materials.⁷

The relationship between grammatical knowledge and reading comprehension has been the focus of extensive empirical research. A meta analysis by Zheng involving over 14,000 students across educational levels found a strong

⁵ Douglas Biber, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. *Longman Grammar of Spoken and Written English*. Pearson Education, 2002.

⁶ Haoyuan Zheng, Xuecong Miao, Yang Dong, and Daniel-Chongbo Yuan, "The relationship between grammatical knowledge and reading comprehension: A meta-analysis," *Frontiers in Psychology* 14 (2023)

⁷ Haoyuan Zheng, Xuecong Miao, Yang Dong, and Daniel-Chongbo Yuan, "The relationship between grammatical knowledge and reading comprehension: A meta-analysis," *Frontiers in Psychology* 14 (2023)

positive correlation between grammatical knowledge and reading comprehension, with the effect size increasing at higher academic stages.⁸ This finding is in line with Srisang and Everatt, who observed that university EFL students with higher grammatical proficiency show better ability to build textual coherence, monitor meaning, and process complex syntactic information.⁹ These skills are necessary for understanding academical discourse, as they allow readers to follow arguments, identify referential ties, and synthesize information from multiple sources.

Research in the context of EFL in Indonesia increasingly emphasizes the importance of grammar for reading comprehension. Grammar mastery is often cited by students as the most challenging aspect of academic reading and writing, yet it is also the most influential factor in their ability to comprehend texts.¹⁰ Prasetya in their quantitative analysis revealed a very strong positive correlation ($r = 0.856$) between grammar mastery and reading comprehension, indicating that explicit grammar instruction and continuous practice are essential for academic success. From a corpus-based perspective, this emphasis on grammar is consistent with Biber's view that grammatical knowledge in academic English consists of recurring patterns in authentic texts, and that awareness of these patterns is essential for constructing meaning during reading.¹¹

⁸ Haoyuan Zheng, Xuecong Miao, Yang Dong, and Daniel-Chongbo Yuan, "The relationship between grammatical knowledge and reading comprehension: A meta-analysis," *Frontiers in Psychology* 14 (2023)

⁹ Pawadee Srisang and John Everatt, "Lower and higher level comprehension skills of undergraduate EFL learners and their reading comprehension," *LEARN Journal: Language Education and Acquisition Research Network* 14, no. 1 (2021): 427–454.

¹⁰ Aldila Fani Prasetya et al., "The correlation between students' grammar mastery and students' reading comprehension," *Journal of English Language and Education* 7, no. 1 (2022): 105–113.

¹¹ Douglas Biber, Susan Conrad, and Randi Reppen, *Corpus Linguistics: Investigating Language Structure and Use* (Cambridge: Cambridge University Press, 1998).

News articles, especially those related to science and technology, are a distinct genre in journalistic writing. News is shaped by a set of news values such as relevance, magnitude, novelty, and the involvement of individuals or elite institutions that guide which events are selected for coverage and how they are presented to the public.^{12 13} These values, as identified by Galtung and Ruge and later refined by Harcup and O'Neill, ensure that news texts are not merely neutral reflections of reality, but constructed narratives that prioritize certain aspects of events over others. The selection process is fundamental to understanding why certain scientific and technological developments receive public attention while others do not, highlighting the importance of critical literacy in the context of EFL.

Science and technology news texts are characterized by high information density, specialized vocabulary, and complex grammatical structures, including nominalization, passive voice, and implied clauses.¹⁴ These linguistic features are not only a reflection of the subject matter but also a result of the news values that prioritize significance and accessibility for a broad audience. For EFL students, especially at the university level, engaging with such texts presents both challenges and opportunities. The authentic language and up-to-date content expose learners to real world discourse, but the complexity of the language can hinder comprehension if students are not equipped with sufficient grammatical and lexical knowledge. This dual nature of science and technology news texts positions them

¹² Harcup, T., & O'Neill, D. (2001). What is News? Galtung and Ruge revisited. *Journalism Studies*, 2(2), 261–280.

¹³ Harcup, T., & O'Neill, D. (2016). What is news? News values revisited (again). *Journalism Studies*, 18(12), 1470–1488.

¹⁴ Harcup, T., & O'Neill, D. (2016). What is news? News values revisited (again). *Journalism Studies*, 18(12), 1470–1488.

as an invaluable resource for developing advanced reading skills and academic literacy in EFL learners.

From a pedagogical perspective, the analysis of science and technology news texts is highly relevant for EFL reading comprehension. Exposure to authentic news discourse helps students develop the ability to process complex grammatical structures, infer meaning from context, and critically evaluate information skills that are essential for academic success and lifelong learning.¹⁵ Furthermore, understanding the news values and linguistic features that shape these texts enables educators to design more effective reading materials and strategies, bridging the gap between classroom instruction and real-world language use. Integrating science and technology news texts into EFL curricula not only enhances linguistic competence but also fosters critical thinking and global awareness among students.

Despite the growing use of corpus based methods in discourse analysis, there remains a notable gap in research that systematically examines grammatical and evaluative patterns in science and technology news texts, especially in relation to EFL reading comprehension at the university level. Most previous studies have focused on either general news genres, online comments, or gender representation in media, often emphasizing lexical frequency, collocation, or evaluative language.^{16 17} While Cavasso & Taboada demonstrate the value of corpus

¹⁵ Harcup, T., & O'Neill, D. (2016). What is news? News values revisited (again). *Journalism Studies*, 18(12), 1470–1488.

¹⁶ Cavasso, L., & Taboada, M. (2021). A corpus analysis of online news comments using the Appraisal framework. *Journal of Corpora and Discourse Studies*, 4, 1-38.

¹⁷ Yuliawati, S., & Ekawati, D. (2023). Sketching women: A corpus analysis of woman representation in the Sundanese Magazine *Manglé* (1958–2019). *Indonesian Journal of Applied Linguistics*, 13(1), 173-182.

annotation for uncovering nuanced evaluative language in online news comments, and Yuliawati & Ekawati show how corpus analysis can reveal shifts in representation and meaning over time, few studies have directly addressed the grammatical patterns that shape comprehension of science and technology news texts for EFL learners.

Furthermore, much of the existing corpus research either focuses on social media language, comments, or gendered representation, and tends to employ either qualitative or quantitative methods in isolation, limiting the generalizability and pedagogical relevance of findings.¹⁸ There is a lack of comparative studies that analyze the alignment (or misalignment) between the grammatical structures found in authentic news texts and those presented in university-level EFL textbooks. This leaves a gap in understanding which grammatical patterns are most prevalent, which are underrepresented in teaching materials, and how these patterns may impact students' ability to comprehend real-world science and technology news.

By addressing this gap, the present study aims to provide a corpus-based analysis of grammatical patterns in science and technology news texts, compare these patterns with those found in university EFL textbooks, and discuss the implications for reading comprehension and curriculum development. This approach will contribute to a more nuanced understanding of the linguistic demands placed on EFL learners and offer practical recommendations for bridging the gap between authentic language use and classroom instruction.

¹⁸ Yuliawati, S., & Ekawati, D. (2023). Sketching women: A corpus analysis of woman representation in the Sundanese Magazine *Manglé* (1958–2019). *Indonesian Journal of Applied Linguistics*, 13(1), 173-182.

1.2. Research Problem

1. What grammatical patterns are most frequently found in contemporary (around December 2025) science and technology news texts written in English?
2. How does the frequency and range of these grammatical patterns compare to those presented in standard university-level EFL textbooks?
3. Which grammatical patterns should be prioritized by EFL learners to develop authentic reading comprehension skills for real-world academic and informational texts?

1.3. Objective of the research

1. To identify and categorize the most frequent grammatical patterns found in contemporary English science and technology news texts.
2. To compare the frequency and distribution of these grammatical patterns with those presented in standard university-level EFL textbooks.
3. To recommend which grammatical patterns should be prioritized in EFL instruction to help students achieve authentic reading proficiency in real-world contexts.

1.4. Significance of the Research

The significance of this research lies in its potential to bridge the gap between authentic English news texts and the grammatical knowledge of EFL learners, which is essential for enhancing reading comprehension skills.

Understanding the specific grammatical patterns that are common in news discourse but less emphasized in EFL textbooks will provide valuable insights for curriculum developers and educators aiming to better prepare students for real-world reading experiences. Additionally, this study can contribute to language teaching practices by identifying grammatical challenges that hinder comprehension, thus guiding more targeted instructional strategies and materials that improve learners' ability to interpret complex informational texts such as news articles.

1.5. Scope and Limitation of the Research

Scope of the Research

This research will focus on analyzing the grammatical patterns present in contemporary English news texts, particularly those in the science and technology domain. The study will compare the frequency and variety of these patterns with those found in English language textbooks used by Indonesian university level EFL students. The research aims to investigate how these grammatical features impact students' reading comprehension abilities within academic contexts.

Limitation of the Research

The study is limited to a selected corpus of science and technology news articles published within the last five years and specific English textbooks commonly used in Indonesian universities. It will primarily emphasize grammatical structures, excluding an in-depth analysis of lexical features. The findings may not be generalizable to other genres of news or to EFL contexts outside of Indonesia.

Additionally, the study will focus on descriptive and correlational analysis without experimental intervention.

1.6. Definition of Key Terms

To ensure clarity and consistency in this study, the following key terms are defined:

a. Corpus Analysis

Corpus analysis is a linguistic research method that involves examining large text collections (corpora) to identify patterns, word frequencies, and language structures.¹⁹ It allows researchers to extract and analyze grammatical patterns in news texts to determine their relevance for university EFL students.²⁰

b. Grammatical Structures

Grammatical structures refer to the rules and patterns that govern the construction of sentences in a language. These include the use of tenses, voice (active/passive), clauses, phrases, and connectors that organize ideas coherently and meaningfully in both spoken and written discourse.

c. Reading Comprehension

Reading comprehension is the ability to understand, interpret, and critically engage with written text. For EFL learners, it involves decoding linguistic

¹⁹ Tony McEnery and Andrew Hardie, *Corpus Linguistics: Method, Theory and Practice* (Cambridge: Cambridge University Press, 2012), 1–2.

²⁰ Douglas Biber et al., *Corpus Linguistics: Investigating Language Structure and Use* (Cambridge: Cambridge University Press, 1998), 4.

structures, grasping meaning from context, and integrating prior knowledge to construct an understanding of texts, such as news articles.

d. News Texts

News text is a genre of journalistic writing that aims to inform readers about recent events, facts, or developments. It is structured to prioritize clarity, accuracy, and relevance, often incorporating specific language features such as impartial tone, factual reporting, and coherence devices to facilitate understanding. News texts typically involve concise sentence constructions, thematic progression, and frequent use of passive voice and nominalizations to efficiently convey complex information.