



A CORPUS LINGUISTICS STUDY OF FREQUENT COLLOCATIONS IN PUBLIC ADMINISTRATION

Misnawati Misnawati¹, Widya Pertiwi Anwar², Zul Astri³

¹Universitas Cahaya Prima

²Doctoral Student at Universitas Negeri Makassar

³Universitas Muslim Maros

Email: misnawati_amir@yahoo.com

ABSTRACT

The study uses a corpus-based analysis to explore the most frequently used word collocations in public administration. The research involved systematically collecting and analyzing texts relevant to public administration, including textbooks, journal articles, and policy documents. Using AntConc software for text analysis, the study identified and categorized 5.883 frequent collocations, providing insights into public administrators' linguistic patterns and communication practices. The findings highlight the importance of specialized vocabulary in public administration and offer a valuable resource for educators and practitioners to enhance their vocabulary knowledge. The study also discusses the pedagogical implications and the potential benefits of incorporating corpus-based word lists into language learning and professional training. This research addresses a significant gap in the literature, emphasizing the need for further investigation into the practical applications of word collocations in public administration.

Keywords: *Public Administration, Corpus-Based Analysis, Word Collocations, Specialized Vocabulary, Linguistic Patterns*

1.1. Introduction

It has been extensively recognized that word lists and corpora are practical tools to help learners improve their vocabulary (Ma & Kelly, 2006; Nation & Waring, 1997; Read, 2000; Schmitt, 1997a). Language plays a crucial role in the teaching process in public administration, where effective communication is essential for policy-making, governance, and public service delivery (Misnawati et al., 2024). Public administration professionals rely on specific terminology and linguistic patterns to convey complex concepts, develop policies, and engage with stakeholders (Gnyawali, 2010; Jordan, 1997). Understanding the most frequent collocations used in this field can provide valuable insights into public administrators' linguistic characteristics and communication practices. Moreover, the linguistic patterns and trends in public administration can provide valuable insights into the evolution of the discipline and the shifting priorities and focus areas within the field (Gray & Jenkins, 1995; Honey, 1967; Welp et al., 2007).

Despite the significant role of language in public administration, there is a lack of comprehensive research on the specific linguistic patterns and frequently used



collocations in this field (Misnawati et al., 2024). Previous studies have focused on the language used in legal, medical, and business contexts, but there has been limited research explicitly targeting the linguistic trends in public administration. Existing studies have often overlooked the unique language requirements and communication practices of public administrators (Masrai et al., 2021; Nagy & Townsend, 2012; Snow, 2010; Szabo et al., 2021), resulting in a limited understanding of the linguistic trends that characterize public administration discourse. Therefore, this research seeks to address this gap by conducting a detailed analysis of the most frequent collocations in public administration, thereby contributing to developing specialized linguistic resources and improving communication effectiveness within this domain.

This research is particularly timely and urgent, given the increasing complexity of governance in the 21st century. With the rise of globalization, digitalization, and new forms of citizen engagement, understanding the evolving language of public administration is crucial for ensuring effective communication, policy implementation, and democratic accountability. By identifying and analyzing critical linguistic trends, this study will provide valuable insights for scholars and practitioners seeking to navigate the changing landscape of public administration and enhance its responsiveness to contemporary challenges. Therefore, the current study aims to explore the most frequently used word collocations in public administration texts, according to a corpus-based analysis, and how the identified frequent collocations reflect the core themes and concepts of public administration.

1.1. Importance of Specialized Language in Public Administration

Public administration relies heavily on precise and effective communication, where specific terminology and linguistic patterns are crucial for policy-making, governance, and public service delivery. Fairclough (2003) highlighted that specialized language helps convey complex concepts clearly and succinctly, facilitating better stakeholder understanding. Public administrators often need to communicate with a diverse audience, including policymakers, government officials, and the general public, making the mastery of relevant vocabulary and collocations essential for their professional efficacy (Honey, 1967).



1.2. Previous Studies on Linguistic Trends in Professional Fields

While extensive research has been conducted on the specialized languages used in legal, medical, and business contexts (Bhatia, 1997; Swales & Swales, 1990), there has been limited focus on the linguistic trends specific to public administration. Gotti's (2003) and Tiersma's (1999) studies have shown how understanding frequent words and collocations in specialized fields can enhance communication effectiveness and professional competence. However, similar comprehensive studies in public administration are notably scarce, creating a gap in the literature this study aims to fill.

1.3. Linguistic Patterns and Collocations in Public Administration

Research on linguistic patterns and collocations involves analyzing how words frequently co-occur, providing insights into common phrases and expressions used within a specific domain (Sinclair, 1991). In public administration, identifying these patterns can reveal the prevalent themes and priorities within the field. For instance, Baker (2006) emphasizes the importance of corpus-based studies in uncovering these linguistic trends, which can inform the development of educational materials, policy documents, and communication strategies tailored to public administration professionals.

1.4. Corpus linguistics

Corpus linguistics, which involves language study through extensive collections of texts, has become a prominent methodology for analyzing linguistic trends (Mcenery & Hardie, 2012). This approach allows researchers to identify the most frequent words and collocations systematically and empirically (Misnawati et al., 2024). Notable corpus linguistics studies, such as those by Biber et al. (1998) and Hyland & Tse (2007), have demonstrated its effectiveness in various fields. Applying this methodology to public administration can provide a robust framework for understanding its specialized language.

1.5. Theoretical Frameworks and Practical Implications

This study's theoretical frameworks include systemic functional linguistics (Teruya & Matthiessen, 2015) and lexical semantics (Cruse, 1986). These frameworks help us understand how language functions within the context of public administration and how



specific word choices and collocations contribute to meaning-making. Practically, the findings from this research can aid in developing linguistic resources such as glossaries, textbooks, and training programs for public administration professionals, enhancing their communication skills and professional competence.

1.2. Research questions

Based on the existing introduction and literature review above, then, the current study proposed research question as follows:

1. What are the most frequently collocations found in public administration texts, according to a corpus-based analysis?
2. What implications do these linguistic trends have for the development of specialized linguistic resources aimed at improving students' communication in public administration?

2. METHOD

The corpus must be carefully selected and organized to create an efficient and fair word list. The researcher collects authentic corpus material from written texts by tracing various sources from books, textbooks, and collections of journal articles that are narrative, general, or academic writings. A step-by-step guide to facilitate this corpus-building process is outlined below (Toriida, 2016):

2.1. *Setting the Target Material*

The first stage involves setting the target material to create a comprehensive and accurate word list for the field of public administration. The selection criteria for this stage focus on identifying texts and materials that are highly relevant to public administration. This includes sources such as textbooks, journal articles, policy documents, lecture notes, and relevant scripts. Textbooks and academic journals are selected for their comprehensive and authoritative coverage of various aspects of public administration. Policy documents, including government reports, white papers, and official statements, are also included to provide insights into practical applications and current trends. Lecture materials and conference papers are also collected in academic and professional training settings. Other relevant writings, such as narrative accounts,



case studies, and graded readings pertinent to public administration, are also incorporated to ensure a well-rounded corpus.

2.2. Collecting the Corpus

The next stage involves collecting the corpus. This begins with a thorough search for physical copies of identified texts in libraries and bookstores. Digital versions of relevant texts are downloaded from academic databases and online repositories. When necessary, permissions are sought to use proprietary materials. Once the materials are collected, they are converted into a digital format suitable for analysis using Anthony's Antword Profiler. For physical documents, high-quality scanning is employed to create readable PDFs or Word documents. It is essential to ensure that all digital versions are accessible and the text is extractable for processing. This stage ensures that the corpus is comprehensive and ready for detailed analysis.

2.3. Eliminating Text

After collecting the corpus, the next stage involves eliminating unnecessary text. The content review process ensures the collected material is relevant to the study's objectives. Text irrelevant or unnecessary to the corpus study is carefully removed, focusing only on content that directly contributes to understanding the vocabulary needs in public administration. Data cleaning is then performed to standardize the format of the text, ensuring consistency across the corpus. This involves deleting off-topic text sections, removing duplicates, and excluding material that does not fit the predefined criteria. Text cleaning tools eliminate formatting issues and non-textual elements, resulting in a clean and relevant corpus for analysis.

2.4. Analyzing the Corpus

The corpus analysis uses text analysis software, specifically AntConc Version 4.2.4 (Anthony, 2022). This software is configured to identify and quantify word frequencies and collocations within the corpus. In addition to software analysis, a manual review is conducted by experts in public administration to ensure accuracy. The findings from the software are cross-referenced with specialized dictionaries and glossaries in public



administration to validate the results. This dual approach ensures that the analysis is both comprehensive and precise.

2.5. Listing the Frequent Collocations

The final stage involves listing the frequent collocations identified in the corpus. Frequency analysis generates a list of collocations that appear most frequently. These frequent items are categorized based on their contextual relevance and importance in public administration. The results are compiled into a detailed report that includes the frequencies and contextual examples of the most frequent collocations. This ensures that the collocations reflect the linguistic trends and terminologies prevalent in public administration discourse, providing valuable insights for professionals and researchers in the field.

3. FINDINGS AND DISCUSSION

3.1. Findings

This study aims to identify the most frequently occurring word lists and word collocations in the field of public administration. The list of words considered to be frequently occurring was recorded by occurring at least 50 times in all the sources or materials analyzed, and words below this criterion were excluded.

Table 1. Number of texts analyzed in the corpus.

References	Number of files	Token	Word types
Website	249		
Books/ Textbooks	59	24.226.582	201,915
Journal articles	253		
	561		

As can be seen in Table 2, 561 references were used to find the word collocation list. After analyzing these references using AntConc software, we found 24.226.582-word tokens and 201.915-word types. In corpus linguistics, "type" and "token" are fundamental concepts that distinguish various aspects of word occurrence in a text or corpus. A "type" refers to a unique instance of words in a corpus. Each distinct form of a word counts as one type, regardless of the number of times it occurs. Meanwhile, a "token" refers to each



instance of a word that appears in the corpus. It represents the total number of words, including repetitions. For example, in the sentence "The fast brown fox jumped over the lazy dog," there are nine-word tokens (given that "the" appears twice) and eight-word types.

When searching using AntConc software, there are 4.058.260 collocations with type 2 words, 12.733.856 with type 3 words, and 18.935.582 collocations with four words with all the number of occurrence frequencies. However, after the review and frequency analysis process, 5.570 collocations of 2 words, 301 collocations of 3 words, and 12 collocations of 4 words were identified. So, the total number of collocations with a frequency of occurrence of 50 times and above amounted to 5.883-word collocations. All of the data can be accessed at <https://www.ycit.or.id/ecopa/collocations.php> for free and online. Examples of collocations with the highest frequency of occurrence are shown in Table 3,4,5 below:

Table 2. Word collocations in 2-words

2-words	Freq.	2-words	Freq.	2-words	Freq.
Public		Relationship			
administration	22221	between	3475	developing countries	2695
Decision making	9383	Civil service	3464	local government	2468
		Public		development	
Public policy	8752	management	3372	administration	2429
		Strategic			
Political science	6957	management	3372	public procurement	2336
				economic	
Public sector	6180	Private sector	3290	development	2275
				sustainable	
University press	5736	Human resources	3091	development	2066
		Resource			
Human resource	4960	management	3060	policy process	1919
		Information			
Public service	4743	systems	2990	economic growth	1890
				competitive	
Policy analysis	4652	Public services	2802	advantage	1882



Policy making	4042	New public	2698	performance management	1781
---------------	------	------------	------	------------------------	------

Table 3. Word collocations in 3-words

3-words	Freq.	3-words	Freq.
Human resource management	2907	Management information systems	1426
New public management	1426	Government printing office	1104
Overall learning experience	1104	Database management system	1103
Subsequent rights restrictions	1103	Civil service reform	1101
Decision-making process	838	Policy making process	838
New public administration	714	Cost-benefit analysis	714
International monetary fund	684	Work-life balance	684
Long term care	657	Human resource planning	657
Self-instructional material	539	Public policy-making	539
Public policy analysis	2907	Decision-Making Processes	350

Table 4. Word collocations in 4-words

4-words	Freq.
Strategic human resource management	198
Management information systems notes	129
Post-new public management	123
Social Affairs population division	112
City County Management Association	88
Civil Service Reform Act	85
Age highest educational attainment	84
Administration notes self-instructional	70
Population age highest educational	56
Government printing office highlights	53
Development self-instructional material	51
Strategic human resource management	198



3.2. Discussion

The results of this corpus analysis show that a large part of word collocations in the field of public administration can be a reference for students and lecturers in the field in the learning process. As described by Xue & Nation (1984), each field needs a specific word list that can meet the needs of native and non-native speakers who have to read and write academic research articles. In addition, many researchers argue that large word lists make practical applications difficult as they can be burdensome for students (Brezina & Gablasova, 2015; Nation & Kyongho, 1995). Therefore, the word list in this study is divided into several sub-lists based on the number of collocation words presented to the students with an alphabetical order display to make it easier for students to learn.

5,883-word collocations created can be helpful for learners in the field of public administration and expand students' vocabulary knowledge to enter the workforce in that field. With these collocations, lecturers can guide students and encourage them to use the word list well. Students will find that the words appearing in the collocation list are made from authentic language commonly used by writers, academics, and practitioners in public administration.

To use and provide the collocation list to students for self-study, lecturers should first consider the students' English competence as the collocation words vary from beginner to advanced level, so choosing a collocation word sub-list that suits the student's needs and level is essential. Thus, Schmitt (2010) clarified that a lecturer can evaluate their students' knowledge before learning the appropriate level of the word list. On the other hand, although using word lists is one of the vocabulary learning strategies mentioned by Schmitt (1997), there are different views on implementing word lists. Some studies suggest that word lists can provide significant benefits and should be integrated into the curriculum (Ma & Kelly, 2006; MacArthur & Littlemore, 2008; Schmitt, 1997, 2010; Vongpumivitch et al., 2009). However, some oppose the use of word lists in the classroom, as the use of word lists is considered one of the least valuable strategies in learning vocabulary and may not be widely accepted in teaching because the language is too difficult for most students. Instead, it has been advocated that vocabulary should be learned from context, not vocabulary-focused learning materials (Boulton, 2009; Fan, 2003; Hulstijn, 2001; Yoon & Hirvela, 2004). For this reason, researchers advocate the



application of corpus-based learning and word lists in improving students' vocabulary knowledge.

3.3. Limitations of the study

The public administration word collocation list is considered to help students' language learning by proposing the practical use of word lists and corpora. To align with pedagogical practice, lecturers should be able to adapt some valuable and practical learning materials as innovations developed from the word list so that the impact of applying the corpus word list can be evaluated. Future research should examine the effects or outcomes of using word lists and assess whether word lists can help students improve their vocabulary knowledge. Some intervening variables may be needed in the practical application of corpus word lists. Students may find learning from the corpus more rewarding if the material is made more meaningful and thus more helpful in their future studies and careers. They may feel more engaged and satisfied as their instructors allow them to study anywhere.

4. CONCLUSIONS

4.1. Conclusion

This study comprehensively analyzes the most frequently used word collocations in public administration, highlighting the critical linguistic patterns prevalent in this field. The research identifies critical collocations that reflect the core themes and concepts of public administration by systematically collecting and analyzing a diverse range of texts, including textbooks, journal articles, policy documents, and other relevant materials. The findings underscore the importance of specialized vocabulary for effective communication in public administration, offering valuable insights for educators and practitioners. The generated word collocation lists serve as a practical resource to enhance the vocabulary knowledge of students and professionals, facilitating better understanding and engagement with public administration discourse. This study contributes significantly to developing tailored linguistic resources, emphasizing the need for further research on the pedagogical impacts and practical applications of corpus-based word lists in language learning and professional training.



5. REFERENCES

- Anthony, L. (2022). *AntConc (Version 4.1.4)*. Waseda University. <https://www.laurenceanthony.net/software>
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. Continuum.
- Bhatia, V. K. (1997). Genre-mixing in academic introductions. *English for Specific Purposes*, 16(3), 181–195.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511804489>
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37–54.
- Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, 36(1), 1–22.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge university press.
- Fairclough, N. (2003). *Analysing discourse* (Vol. 270). routledge London.
- Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, 87(2), 222–241.
- Gnyawali, B. D. (2010). *The need of English in public administration* [Thesis, Central Department of Education]. <https://elibrary.tucl.edu.np/handle/123456789/21216>
- Gotti, M. (2003). *Specialized discourse: Linguistic features and changing conventions*. P. Lang.
- Gray, A. T., & Jenkins, B. (1995). From public administration to public management: reassessing a revolution? *Public Administration*, 73(1), 75–99. <https://doi.org/10.1111/j.1467-9299.1995.tb00818.x>
- Honey, J. C. (1967). A report: Higher education for public service. *Public Administration Review*, 27(4), 294. <https://doi.org/10.2307/973346>
- Hulstijn, J. H. (2001). *Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity*.
- Hyland, K., & Tse, P. (2007). Is There an “Academic Vocabulary”? *TESOL Quarterly*, 41(2), 235–253. <https://doi.org/10.1002/j.1545-7249.2007.tb00058.x>
- Jordan, R. R. (1997). *English for academic purposes: A guide and resource book for teachers*. Cambridge University Press.
- Ma, Q., & Kelly, P. (2006). Computer assisted vocabulary learning: Design and evaluation. *Computer Assisted Language Learning*, 19(1), 15–45.



- MacArthur, F., & Littlemore, J. (2008). A discovery approach to figurative language learning with the use of corpora. *APPLICATIONS OF COGNITIVE LINGUISTICS*, 6, 159.
- Masrai, A., Milton, J., El-Dakhs, D. A. S., & Elmenshawy, H. (2021). Measuring the contribution of specialist vocabulary knowledge to academic achievement: disentangling effects of multiple types of word knowledge. *Asian-Pacific Journal of Second and Foreign Language Education*, 6, 1–15.
- Mcenery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.
- Misnawati, M., Atmowardoyo, H., Sulaiman, I., Yusriadi, Y., & Rahman, A. (2024). Unveiling the lecturers' and students' needs in English for public administration program: Essential vocabulary topics, instructional methods, and learning challenges. *REGISTER JOURNAL*, 17(1), 100–122. <https://doi.org/10.18326/rgt.v17i1.100-122>
- Misnawati, M., Nur, S., & Tahir, S. Z. (2024). Corpus linguistics today: A qualitative approach. *Research and Innovation in Applied Linguistics [RIAL]*, 2(1), 45–62. <https://doi.org/10.31963/rial.v2i1.4486>
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91–108. <https://doi.org/10.1002/RRQ.011>
- Nation, P., & Kyongho, H. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, 23(1), 35–41.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. *Vocabulary: Description, Acquisition and Pedagogy*, 14(1), 6–19.
- Read, J. A. S. (2000). *Assessing vocabulary*. Cambridge university press.
- Schmitt, N. (1997a). Vocabulary learning strategies. *Vocabulary: Description, Acquisition and Pedagogy*, 199227, 4–6.
- Schmitt, N. (1997b). Vocabulary learning strategies. *Vocabulary: Description, Acquisition and Pedagogy*, 199227, 4–6.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation* (Describing). Oxford University Press.
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328(5977), 450–452. <https://doi.org/10.1126/science.1182597>
- Swales, J. M., & Swales, J. (1990). *Genre analysis*. Cambridge university press.
- Szabo, C. Z., Stickler, U., & Adinolfi, L. (2021). Predicting the academic achievement of multilingual students of English through vocabulary testing. *International Journal of Bilingual Education and Bilingualism*, 24(10), 1531–1542.



- Teruya, K., & Matthiessen, C. (2015). Halliday in relation to language comparison and typology. *The Bloomsbury Companion to MAK Halliday*, 427–452.
- Tiersma, P. M. (1999). *Legal language*. University of Chicago Press.
- Toriida, M.-C. (2016). Steps for creating a specialized corpus and developing an annotated frequency-based vocabulary list. *TESL Canada Journal/Revue TESL Du Canada*, 34(11), 87–105. <https://doi.org/http://dx.doi.org/1018806/tesl.v34i1.1255>
- Vongpumivitch, V., Huang, J., & Chang, Y.-C. (2009). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes*, 28(1), 33–41.
- Welp, Y., Urgell, F., & Aibar, E. (2007). From bureaucratic administration to network administration? An empirical study on e-government focus on Catalonia. *Public Organization Review*, 7(4), 299–316. <https://doi.org/10.1007/s11115-007-0038-z>
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215–229.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257–283.