

THESIS

**LEXICAL BUNDLES IN MULTIPLE DISCIPLINES: A
CORPUS-BASED STUDY IN JOURNAL ARTICLES**

PUBLISHED BY PLOS ONE



NAME: MOHAMAD IRHAM POLUWA

STUDENT NO: 121914253004

**MASTER OF LINGUISTICS STUDY PROGRAM
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA**

2022



THESIS

**LEXICAL BUNDLES IN MULTIPLE DISCIPLINES: A
CORPUS-BASED STUDY IN JOURNAL ARTICLES**

PUBLISHED BY PLOS ONE



NAME: MOHAMAD IRHAM POLUWA

STUDENT NO: 121914253004

**MASTER OF LINGUISTICS STUDY PROGRAM
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA**

2022

LEXICAL BUNDLES IN MULTIPLE DISCIPLINES: A CORPUS-BASED
STUDY IN JOURNAL ARTICLES PUBLISHED BY PLOS ONE

THESIS

Submitted in Partial Fulfillment of the Requirements for obtaining Master's Degree
in Master's Program in Linguistics, Faculty of Humanities, Universitas Airlangga

By
MOHAMAD IRHAM POLUWA
121914253004

MASTER OF LINGUISTICS STUDY PROGRAM
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
2022

THESIS SUBMISSION APPROVAL FORM

This is to certify that the Master's thesis of
Title: Lexical Bundles in Multiple Disciplines: A Corpus-Based Study in Journal
Articles Published by PLOS One

Name: Mohamad Irham Poluwa
Student Number: 121914253004

has been approved to be examined

Surabaya, June 10, 2022

Advisor I



Dr. Ni Wayan Sartini, Drs., M.Hum.
NIP 196308111990022001

Advisor II



Nurul Fitri Hapsari, S.S., M.A.
NIP 198307142012122002

Head of Study Program



Dr. Ni Wayan Sartini, Drs., M.Hum.
NIP 196308111990022001

MASTER OF LINGUISTICS STUDY PROGRAM
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
2022

This thesis has been approved to be examined and accepted by the Board
Examiners of Magister of Linguistics Study Program, Faculty of Humanities,
Universitas Airlangga

Surabaya, June 20, 2022

The Head of Examiner



Dr. Hariawan Adji, Drs., S.T., M.Kes.
NIP 196911101993031004

Examiner I



Dr. Ni Wayan Sartini, Drs., M.Hum.
NIP 196308111990022001

Examiner II



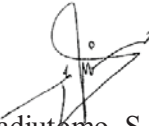
Nurul Fitri Hapsari, S.S., M.A.
NIP 198307142012122002

Examiner III



David Segoh, S.S., M.Ed.
NIP 197304152007011001

Examiner IV



Dwi Anggoro Hadiutomo, S.S., M.Hum., Ph.D.
NIP 197312052002121001

MASTER OF LINGUISTICS STUDY PROGRAM
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
2022

ACKNOWLEDGEMENTS

All praise is due to Allah SWT, the lord of the worlds, the king of the kings, the almighty who has given me knowledge, power, and opportunities, so that I can finally complete my thesis. Peace and Salutations always be given to the prophet Muhammad SAW, as the messenger of Allah who has changed human civilization from the darkened era of stupidity and ignorance into the enlightenment era of knowledge and wisdom.

First of all, I would like to express my gratitude to my first advisor, Mrs. Dr. Ni Wayan Sartini, Drs., M.Hum., who has helped and motivated me with her insightful feedbacks and suggestions towards my journey to complete this thesis. I would also like to express my gratitude to my current second advisor, Mrs. Nurul Fitri Hapsari, S.S., M.A, for giving me constructive feedbacks to the study that I have conducted, so that I can pass all of obstacles and hindrances that I need to overcome in finishing this thesis.

As for my former second advisor, Mr. Totok Suhardijanto, M.Hum., Ph.D., I will never forget all the good things that you have shared with me. Thank you for being such a good example on how an educator, especially a lecturer, must behave. I know my gratitude will not be enough to thank you for all knowledge, insights, experiences, and memories that you have shared with me either in the corpus linguistics class or in several occasions of our thesis consultations. May you rest in peace, sir.

Second, I would like to express my sincere gratitude to my family, especially my parents and brothers for all the good prayers and moral supports. I know this thesis will never be completed without all these people beside me. As for my beloved wife, Nafilaturif'ah, thank you for being such a good wife. Thank you for accompanying me to walk on the same path as a longlife learner. Thank you for becoming my ray of hope, so that I can always remember the reason why I keep fighting. It is you who brought meaning into my life and completing all my shortcomings to become a better person, a better learner, a better man, and a better husband for you.

Third, I would like to express my gratitude to all teachers and lectures who have taught and shared all their knowledge, experiences, insights, and all precious things that I need to know in order to survive and encounter all obstacles that I need to overcome in this eternal journey of life. I would also like to express my gratitude to all my classmates in the master's program in Linguistics for all good memories that we have shared together. Last but not least, I would like to express my gratitude to Fadil, Dhimas, Pandu, Alfian, Adrian, Abi for still being a good friend even after all this time, and also to all of my colleagues and friends who I cannot say one by one here.

Surabaya, 9 July 2022

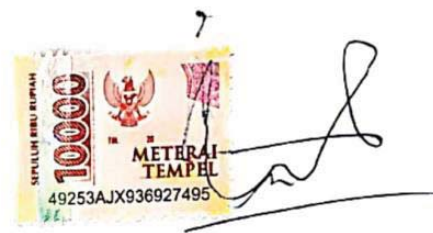
Mohamad Irham Poluwa

DECLARATION OF AUTHORSHIP

I, the undersigned, declare that:

1. This research is original research that has never been used for getting a master's degree in Universitas Airlangga or other universities;
2. This research is made by myself without help from another party except for guidance from my supervisors;
3. This research is not plagiarism from another research published by other people, unless in writing it is clearly stated as a reference in the manuscript with the author's name mentioned and included in the reference page;
4. At a later time, it is found that this thesis is a product of plagiarism, I will accept any consequences that may be imposed on me by the university.

Surabaya, 8 July 2022

A yellow revenue stamp from Indonesia, featuring a Garuda emblem and the text '1000', 'METERAI TEMPEL', and '49253AJX936927495'. A handwritten signature is written over the stamp.

Mohamad Irham Poluwa
NIM. 121914253004

ABSTRACT

LEXICAL BUNDLES IN MULTIPLE DISCIPLINES: A CORPUS-BASED STUDY IN JOURNAL ARTICLES PUBLISHED BY PLOS ONE

Several studies have shown that lexical bundles differ among multiple disciplines. The present study investigates the distribution of lexical bundles among multiple disciplines (i.e., Chemistry, Mathematics, Psychology, and Sociology) based on journal articles published in *PLOS ONE* journal. The present study is not only focused on the analysis of structural forms and functional classifications but is also concerned with the analysis of semantic preference as an attempt to extend the analysis of functional classifications. By applying the corpus-based approach and qualitative approach, the present study is conducted based on the bottom-up model of research. This study indicates that the variation of lexical bundles among the multiple disciplines is relatively affected by the characteristics of each discipline which is rooted in their respective major areas. Besides, it is also found that the semantic preferences of these multiple disciplines are similar and reflect the characteristics of journal articles in general.

Keywords: lexical bundles, structural forms, functional classifications, semantic preference

ABSTRAK

SIMPUL LEKSIKAL DALAM BEBERAPA DISIPLIN ILMU: SEBUAH PENELITIAN BERBASIS KORPUS DALAM JURNAL ARTIKEL YANG DIPUBLIKASIKAN OLEH PLOS ONE

Beberapa penelitian telah menunjukkan bahwa simpul leksikal berbeda dalam beberapa disiplin ilmu. Penelitian ini menganalisis persebaran dari simpul leksikal dalam beberapa disiplin ilmu seperti Kimia, Matematika, Psikologi dan Sosiologi berdasarkan artikel ilmiah yang dipublikasikan oleh jurnal *PLOS ONE*. Tidak hanya berfokus pada bentuk struktural dan klasifikasi fungsional, penelitian ini juga berfokus pada preferensi semantik sebagai upaya perluasan lingkup analisis dari klasifikasi fungsional. Dengan menerapkan pendekatan penelitian berbasis korpus dan pendekatan kualitatif, penelitian ini dilakukan berdasarkan model analisis dari bawah ke atas (*bottom-up*). Penelitian ini menunjukkan bahwa variasi simpul leksikal dalam beberapa disiplin ilmu yang digunakan dalam penelitian ini, dipengaruhi oleh karakteristik dari masing-masing disiplin ilmu berdasarkan pada rumpun keilmuan yang lebih luas dari masing-masing disiplin ilmu tersebut. Selain itu, penelitian ini juga menunjukkan bahwa disiplin ilmu yang digunakan dalam penelitian ini memiliki preferensi semantik yang sama serta sesuai dengan karakteristik dari artikel ilmiah secara umum.

Kata kunci: *simpul leksikal, bentuk struktural, klasifikasi fungsional, preferensi semantik*

TABLE OF CONTENTS

INSIDE COVER PAGE.....	i
INSIDE TITLE PAGE	ii
THESIS SUPERVISORS' APPROVAL FORM.....	iii
THESIS EXAMINERS' APPROVAL FORM	iv
ACKNOWLEDGEMENTS.....	v
DECLARATION OF AUTHORSHIP	vii
ABSTRACT.....	viii
ABSTRAK	ix
TABLE OF CONTENTS	x
LIST OF TABLES.....	xiii
LIST OF FIGURES	xv
CHAPTER 1 INTRODUCTION	1
1.1 Background of the Study	1
1.2 Research Questions	6
1.3 Objectives of the Study	7
1.4 Significance of the Study.....	7
CHAPTER II LITERATURE REVIEW	9
2.1 Theoretical Framework	9
2.1.1 Corpus Linguistics.....	9
2.1.1.1 Concordance	12
2.1.2 Lexical Bundles	12
2.1.2.1 Structural Forms of Lexical Bundles.....	15

2.1.2.2 Functional Classifications of Lexical Bundles	15
2.1.3. Semantic Preference	16
2.2 Related Studies	19
CHAPTER III RESEARCH METHOD	32
3.1 Research Approach	32
3.2 Source of Data.....	33
3.3 Method of Data Collection.....	34
3.4 Method of Data Analysis.....	39
CHAPTER IV RESULTS AND DISCUSSION	46
4.1 Structural Forms of Lexical Bundles in the Four Corpora	46
4.1.1 Structural Forms of Lexical Bundles in CC.....	47
4.1.2 Structural Forms of Lexical Bundles in MC.....	50
4.1.3 Structural Forms of Lexical Bundles in PC	53
4.1.4 Structural forms of Lexical Bundles in SC.....	56
4.2 Functional Classifications of Lexical Bundles in the Four Corpora.....	64
4.2.1 Functional Classifications of Lexical Bundles in CC	64
4.2.2 Functional Classifications of Lexical Bundles in MC.....	68
4.2.3 Functional Classifications of Lexical Bundles in PC.....	71
4.2.4 Functional Classifications of Lexical Bundles in SC	74
4.3 Semantic Preference of the Shared Bundle in the four Corpora	86
4.3.1 Topic Introduction/Focus	88
4.3.2 Topic Elaboration/Clarification.....	89
CHAPTER V CONCLUSION AND SUGGESTIONS	96

5.1 Conclusion	96
5.2 Suggestions	98
References	100
Appendices.....	109
Appendix 1	109
Appendix 2	115
Appendix 3	124

LIST OF TABLES

Table 3.1 Description of Corpora of Each Discipline.....	39
Table 4.1 Structural Forms of Lexical Bundles in CC.....	48
Table 4.2 Structural Forms of Lexical Bundles in MC	51
Table 4.3 Structural Forms of Lexical Bundles in PC	53
Table 4.4 Structural Forms of Lexical Bundles in SC	56
Table 4.5 Distribution of Major Categories of Lexical Bundles in the Four Corpora	60
Table 4.6 The Comparison of the Noun Phrase-based in the Four Corpora	63
Table 4.7 Functional Classifications of Lexical Bundles in CC	65
Table 4.8 Functional Classifications of Lexical Bundles in MC.....	68
Table 4.9 Functional Classifications of Lexical Bundles in PC	71
Table 4.10 Functional Classifications of Lexical Bundles in SC	74
Table 4.11 Distribution of Major Categories of Functional Classifications in the Four Corpora	78
Table 4.12 The Shared Bundles in Pair-up Corpus.....	85
Table 4.13 The Use of the Bundle “in the present study” based on Discourse Organizers.....	87
Table 4.14 The Collocations of “in the present study” in Topic Introduction/Focus.....	88
Table 4.15 The Collocations of “in the present study” in Description/Elaboration	90

Table 4.16 The Collocations of “in the present study” in Clarification/Opinion	
.....	92

LIST OF FIGURES

Figure 3.1 AntCorGen Software Version 1.1.2	35
Figure 3.2 Processing the Lexical Bundles Using AntGram Software Version 1.3.0.....	41
Figure 3.3 Processing Concordances of The Lexical Bundles Using AntConc Software Version 3.5.9.....	44
Figure 3.4 The File View of the Lexical Bundles Processed by AntConc Software Version 3.5.9.....	45