

On Second Language Learner Acquisition of English Collocations

James Martin Rogers

Meijo University, Japan

Series in Education



VERNON PRESS

Copyright © 2024 Vernon Press, an imprint of Vernon Art and Science Inc, on behalf of the author.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Vernon Art and Science Inc.

www.vernonpress.com

In the Americas:
Vernon Press
1000 N West Street, Suite 1200
Wilmington, Delaware, 19801
United States

In the rest of the world:
Vernon Press
C/Sancti Espiritu 17,
Malaga, 29006
Spain

Series in Education

Library of Congress Control Number: 2023947657

ISBN: 978-1-64889-797-9

Cover design by Vernon Press. Cover image by James Martin Rogers.

The supplementary material can be downloaded from <https://vernonpress.com/book/1877>

Product and company names mentioned in this work are the trademarks of their respective owners. While every care has been taken in preparing this work, neither the authors nor Vernon Art and Science Inc. may be held responsible for any loss or damage caused or alleged to be caused directly or indirectly by the information contained in it.

Every effort has been made to trace all copyright holders, but if any have been inadvertently overlooked the publisher will be pleased to include any necessary credits in any subsequent reprint or edition.

For my daughter

Table of Contents

| | | |
|-----------|---|------|
| | List of Figures and Tables | vii |
| | Preface | xi |
| | Acknowledgments | xiii |
| Chapter 1 | Introduction | 1 |
| Chapter 2 | Review of the Literature | 5 |
| | Introduction | |
| | Overview of Second Language Learning Research | |
| | On Learners' Lack of Collocational Fluency | |
| | On the Learning Burden of Collocations | |
| | On Approaching and Defining Collocations | |
| | What is a Collocation? | |
| | Types vs. Lemmas vs. Word Families | |
| | On Semantic Transparency | |
| | On Concrgramming, MWU Length, and Colligation | |
| | On the Value of Collocations | |
| | On the Lack of Research and Resources | |
| | On the Direct Teaching of Collocations | |
| | On Utilizing Corpus Data to Identify Collocations | |
| | On Using L1-L2 Congruency to Identify Useful Collocations | |
| | Conclusion | |
| Chapter 3 | Addressing Gaps in the Research | 37 |
| | Introduction | |
| | The Gaps in the Research | |

Research Methods and Techniques

Research Paradigm

Data Source, Collection Methodology, and Analysis

Filling the Gaps in the Research

What is an Ideal Corpus Frequency Data Cut-off for Identifying High-frequency General English MWUs?

Is Corpus Dispersion Data Reliable for Identifying High-frequency General English MWUs?

Is Corpus Chronological Data Reliable for Identifying High-frequency General English MWUs?

Is Consideration for Colligation an Important Criterion for Identifying High-frequency General English MWUs?

What Percentage of High-frequency General English MWUs is Deemed by Fluent Speakers Worthy of Expanding Beyond Their Most Frequent Exemplar?

What Percentage of MWUs Most Representative of High-frequency General English Lemmatized Congrams Has Low Semantic Transparency?

What Percentage of MWUs Most Representative of High frequency General English Lemmatized Congrams Has Low L1-L2 Congruency With Japanese, Chinese, and Korean?

Is Fluent Speaker Intuition Reliable Regarding High-frequency Vocabulary Usage in Context Creation?

What is Japanese University Students' Knowledge of MWUs Most Representative of High-frequency General English Lemmatized Congrams?

Conclusion

| | | |
|-----------|--------------------------------------|-----|
| Chapter 4 | Implications and Applications | 87 |
| Chapter 5 | Conclusion | 101 |
| | References | 105 |
| | Index | 115 |

List of Figures and Tables

List of Figures

| | | |
|--------------------|---|----|
| Figure 3.1. | Percentage of Items Accurately and Erroneously Flagged for Balanced Dispersion Data Distribution at All Three Parameters | 48 |
| Figure 3.2. | Total Items Erroneously Flagged or Judged Unbalanced Which Were Not Flagged | 48 |
| Figure 3.3. | Percentage of Items Accurately and Erroneously Flagged for Balanced Chronological Data Distribution at all Three Parameters | 55 |

List of Tables

| | | |
|-------------------|--|----|
| Table 2.1. | The Three Approaches' Ability to Identify Common Collocates of the Verb 'Play' | 13 |
| Table 2.2. | High-frequency Collocations for the Four Most Frequent Words in the Word Family for 'Govern' According to the COCA (Top Frequencies in Bold) | 16 |
| Table 2.3. | A Sample of Data From the COCA for a Congram Search for the Lemmas 'Provide' and 'Support' | 19 |
| Table 2.4. | Top Three MWUs for the Lemmas 'Provide' and 'Support' Found After Examining 500 Concordance Strings in the COCA | 20 |
| Table 2.5. | MWUs Identified From 500 Example Sentences in Which the Lemmas 'Come' and 'Term' Both Occur in | 21 |
| Table 2.6. | A Comparison Between Two MWU Searches, One With and One Without Consideration for a Specific Type of Colligation | 22 |
| Table 3.1. | Word Frequency Breakdown of Lemma Pairs Occurring Once Per Million Tokens According to Vocabprofile's 25,000-word Families of the BNC and COCA | 44 |
| Table 3.2. | Criteria for Rating the Value of Collocates for Learners of General English | 47 |
| Table 3.3. | Items Found Not to be Worthy of Inclusion Because They Were Either Inappropriate Language, Grammatical Formulations, Duplicates, or Compound Nouns | 49 |

| | | |
|--------------------|--|----|
| Table 3.4. | Samples of Items Flagged for Having Unbalanced Dispersion in Each of the Four Most Common Genres at All Three Parameters | 49 |
| Table 3.5. | A Sample of Pairs Flagged for Having Unbalanced Dispersion at All Three Parameters Judged to be Erroneously Flagged by a Fluent Speaker | 51 |
| Table 3.6. | A Sample of Pairs Judged to be of Little Use to General Learners Not Flagged for Having Unbalanced Dispersion by Any of the Three Parameters | 52 |
| Table 3.7. | Samples of Items Accurately Flagged at All Three Parameters (2.5, 5, and 10) and Items Judged to Have Chronological Issues Not Flagged by Any of the Parameters (X) | 56 |
| Table 3.8. | The Top MWU Identified When 500 and 1,000 Example Sentences Were Utilized | 59 |
| Table 3.9. | Amount of Top MWUs That Were Affected by Each of the Colligational Treatments | 61 |
| Table 3.10. | Comparison Between Ten Random Samples of Top MWUs Affected by the Colligational Treatment for Pre-nominal Possessive Pronouns and the Results That Would Have Occurred Without the Treatment | 62 |
| Table 3.11. | Comparison Between Ten Samples of Top MWUs Affected by the Colligational Treatment for Subject Pronouns and the Results That Would Have Occurred Without the Treatment | 63 |
| Table 3.12. | Comparison Between Ten Random Samples of Top MWUs Affected by the Colligational Treatment for Cardinal Numbers and the Results That Would Have Occurred Without the Treatment | 64 |
| Table 3.13. | MWUs Identified From 500 Example Sentences in Which the Lemma Pair ‘Come’ and ‘Term’ Both Occur | 67 |
| Table 3.14. | Semantic Transparency Ratings of the Collocations (Percentage of Total Items in Parenthesis) | 69 |
| Table 3.15. | L1-L2 Congruency Ratings of High-frequency English MWUs With Japanese Translations | 71 |
| Table 3.16. | L1-L2 Congruency Ratings of High-frequency English MWUs With Chinese Translations | 71 |
| Table 3.17. | L1-L2 Congruency Ratings of High-frequency English MWUs With Korean Translations | 72 |
| Table 3.18. | Word Family Frequency Breakdown of Formulaic Phrases Using RANGE | 74 |

| | | |
|--------------------|--|----|
| Table 3.19. | Word Family Frequency Breakdown of Formulaic Phrases Within Example Sentences Created Using Fluent Speaker Intuition Using RANGE | 75 |
| Table 3.20. | Vocabprofiler Breakdown of Top 3,000 Word Family Words Not Covered by Example Sentences Created Using Fluent Speaker Intuition | 77 |
| Table 3.21. | L1-L2 Congruency Ratings of MWUs Selected for Testing Students' Collocational Fluency | 79 |
| Table 3.22. | Mean Scores for Test Items Organized by Frequency Level | 81 |
| Table 3.23. | Multiple Regression Analysis and Correlation Coefficient with TOEFL as the Dependent Variable and Item Frequency as the Independent Variable | 81 |
| Table 3.24. | Multiple Regression Analysis and Correlation Coefficient with TOEFL Score as the Dependent Variable and L1-L2 Congruency as the Independent Variable | 82 |
| Table 4.1. | Top Ten TOEIC Gains in Comparison to Minutes Studied with the App | 98 |

Preface

Gaining control of collocations is a very important part of language learning, but learners are typically slow in sounding native-like in their use of collocations. One of the reasons for this is that the majority of collocations occur infrequently. It is useful to bear in mind that the frequency of any collocation will be lower than the frequency of its lowest-frequency member.

Collocations have the chance of being learned across the four strands of a well-balanced course. In the meaning-focused input strand of a course, collocations will be met through listening and reading. Because learning through the strand of meaning-focused input is largely incidental learning, large quantities of input through extensive reading and extensive listening are needed to get enough repetitions of particular collocations for them to stay in memory. Collocations also have the chance to be learned through meaning-focused output, that is, through speaking and writing, where learners have the opportunity to turn receptive knowledge into productive use and to strengthen and enrich their knowledge of collocations. The fluency development strand is likely to be a very important strand in the development of collocational knowledge as the pressure to use language faster encourages a focus on restructuring knowledge to achieve fluency by moving from the processing of single words to greater use of multiword units. These three meaning-focused strands all depend on incidental learning and frequency of occurrence.

The strand of language-focused learning provides the opportunity to speed up the learning of collocations through deliberate teaching and deliberate learning. Such deliberate teaching and learning are most efficiently done if there are well-researched lists of collocations to draw on. This book reports on research that has been done with this goal in mind.

The case for giving deliberate attention to collocations is strengthened by research that suggests that, as with the deliberate learning of vocabulary, the deliberate learning of collocations simultaneously results in both implicit knowledge and explicit knowledge. That is, the deliberate learning of collocations will provide the kind of knowledge that is needed for normal language use.

In order for something to be learned, there are three requirements of attention. Firstly, what needs to be learned needs to be focused on. That is, the form, the meaning, and/or the use of the collocation needs to be focused on. Secondly, there needs to be a quantity of attention. In other words, there

needs to be repeated encounters and good attention for each encounter. The greater the amount of attention, the more likely the collocation is to be remembered. Thirdly, there needs to be good quality attention. In the case of collocation, the quality of attention can be increased by looking at how the parts of the collocation relate to the meaning of the whole, the use of the collocation in a variety of contexts, and visualizing the meaning of the collocation. These requirements of attention apply to both incidental and deliberate learning. In incidental learning, the three requirements are typically not as strongly applied as in deliberate learning. In incidental learning of collocations, as in listening and reading, there tends to be brief attention to the form, with similarly brief meaning recall. Each incidental meeting tends to involve a small amount of attention, but with repeated meetings, there can be a cumulative effect. The quality of attention is typically not very deep, involving a brief focus on the form and contextual meaning.

The research in this book provides very useful support for the deliberate learning of collocations. It makes sure that, in deliberate learning, attention is focused on the most useful items to learn.

Paul Nation

Acknowledgments

I would like to thank all research assistants and translators who volunteered their time for my research projects in the past. Without their help, I wouldn't have been able to tackle the large-scale research projects this book will describe. Filling the gaps in the research would have also been difficult to accomplish if not for the software solution that was provided by Dr. Laurence Anthony.

PAGES MISSING
FROM THIS FREE SAMPLE

References

- Ackermann, K., & Chen, Y. (2013). Developing the academic collocation list (ACL) - A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, 12(4), 235-247.
- Aghbar, A. (1990, October). *Fixed expressions in written texts: Implications for assessing writing sophistication* [Paper presentation]. Meeting of the English Association of Pennsylvania State System Universities, United States.
- Al-Zahrani, M. (1998). *Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi university* [Unpublished doctoral dissertation]. University of Pennsylvania.
- Alexander, R. (1984). Fixed expressions in English: Reference books and the teacher. *ELT Journal*, 38(2), 127-134.
- Almela, M., & Sanchez, A. (2007). Words as "lexical units" in learning/teaching vocabulary. *International Journal of English Studies*, 7(2), 21-40.
- Anthony, L. (2013). *AntWordPairs (Version 1.0.2)* [Computer software]. Waseda University. <http://www.antlab.sci.waseda.ac.jp/>
- Anthony, L. (2018). *AntConc (Version 3.5.8)* [Computer software]. Waseda University. <http://www.antlab.sci.waseda.ac.jp/>
- Arabski, J. (1979). *Errors as indicators of the development of interlanguage*. Uniwersytet Slaski.
- Avery, P., & Baker, J. (1997). Mapping learning at the secondary level. *Clearing House*, 70(5), 279-285.
- Bahns, J., & Eldaw, M. (1993). Should we teach ESL students collocations? *System*, 21(1), 101-114.
- Barghamadi, Rogers, J., Arciuli, J., Han, W., & Muller A. (2023). L1-L2 congruency as a criterion to identify collocations based on contrastive analysis. *Australian Journal of Applied Linguistics*, 6(1), 1-14.
- Bartan, O. S. (2019). Lexical collocation errors in literary translation. *Dil Dergisi*, 1(170), 73-88.
- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education.
- Biskup, D. (1992). L1 influence on learners' renderings of English collocations: A Polish/German empirical study. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 1-12). Macmillan.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10, 245-261.

- Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 23, 321-343.
- Bolinger, D. (1968). *Aspects of language*. Harcourt Brace Jovanovich, Inc.
- Bolinger, D. (1976). Meaning and memory. *Forum Linguisticum*, 1(1), 1-14.
- British National Corpus (1994). <http://www.natcorp.ox.ac.uk/corpus/index.xml>
- Byrd, P., & Coxhead, A. (2010). 'On the other hand': Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5, 31-64.
- Channell, J. (1981). Applying semantic theory to vocabulary teaching. *English Language Teaching Journal*, 35, 115-122.
- Chan, T., & Liou, H. (2005). Online verb-noun collocation instruction with the support of a bilingual concordancer. *Selected Papers From the Fourteenth International Symposium on English Teaching* (pp. 270-281). Crane Publishing Co., Ltd.
- Chen, P. (2002). *A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan* [Unpublished master's thesis]. National Taiwan Normal University.
- Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics*, 11(4), 411-433.
- Chomsky, N. (1957). *Syntactic structures*. Mouton and Company.
- Chon, Y., & Shin, D. (2009). Collocations in L2 writing and rater's perceived writing proficiency. *Korean Journal of Applied Linguistics*, 25(1), 101-129.
- Cobb, T. (2013). *VocabProfiler VP Compleat BNC-COCA-25*. <https://www.lex tutor.ca/vp/comp/>
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29, 72-89.
- Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 70-86). Language Teaching Publications.
- Cowan, L. (1989). *Towards a definition of collocation* [Unpublished master's thesis]. Concordia University.
- Cowie, A. (1978). The place of illustrative material and collocations in the design of a learner's dictionary. In P. Strevens (Ed.), *In honour of A.S. Hornby* (pp. 127-139). Oxford University Press.
- Cowie, A. (Ed.). (1998). *Phraseology: theory, analysis, and applications*. Oxford University Press.
- Davies, M. (2008). *The corpus of contemporary American English: 425 million words, 1990-present*. <http://corpus.byu.edu/coca/>
- Davies, M. (2010). *Word list plus collocates*. <http://www.wordfrequency.info/purchase1.asp?i=c5a>
- Daulton, F. (2008). *Japan's built-in lexicon of English-based loanwords*. Multilingual Matters Ltd.
- Davoudi, M., & Behshad, A. (2015). Collocational use: a contrastive analysis of strategies used by Iranian EFL learners. *Theory and Practice in Language Studies*, 5(12), 2646-2652.

- Dechert, H. (1983). How a story is done in a second language. In C. Faerch & G. Kasper, (Eds.), *Strategies in interlanguage communication* (pp. 175-195). Longman.
- Dechert, H., & Lennon, P. (1989). Collocational blends of advanced second language learners: A preliminary analysis. In W. Olesky (Ed.), *Contrastive pragmatic* (pp. 131-168). Benjamins.
- DeCock, S., Granger, S., Leech, G. and McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. In S. Granger (Ed.), *Learner English on computer* (pp. 67-79). Longman.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-262). CUP.
- Du, L., Elgort, I., & Siyanova-Chanturia, A. (2021). Cross-language influences in the processing of multiword expressions: From a first language to second and back. *Frontiers in Psychology, 12*(666520), 1-18.
- Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *Journal of English for Specific Purposes, 28*(3), 157-179.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics, 47*, 157-177.
- Ellis, N. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33-68). Cambridge University Press.
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.
- Eriksson, A. (2012). Pedagogical perspectives on bundles: Teaching bundles to doctoral students in biochemistry. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 195-211). Masaryk University Press.
- Fayez-Hussein, R. (1990) Collocations: The missing link in vocabulary acquisition amongst EFL learners. In J. Fisiak (Ed.), *Papers and studies in contrastive linguistic: The Polish English contrastive project, 26* (pp. 123-126). Adam Mickiewicz University.
- Firth, J. (1957). A synopsis of linguistic theory. 1930-1955. In F. Palmer (Ed.), *Selected papers of J.R. Firth 1952-59* (pp. 168-205). Longman.
- Francis, G. (1993). A corpus-driven approach to grammar: Principles, methods and examples. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 137-156). John Benjamins.
- Furukawa, J., Ford, B., Ayson, E., Cambra, K., Takahashi, L., & Yoshina, K. (1998, January 17). *Effects of a cognitive processing strategy on spelling, definitions, and reading* [Conference presentation]. The 20th Annual Conference of the Hawaii Educational Research Association. Honolulu, HI, United States.
- Gairns, R., & Redman, S. (1986). *Working with words. A guide to teaching and learning vocabulary*. Cambridge University Press.

- Gasser, M. (1990). Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition*, 12, 179-199.
- Gitsaki, C. (1996). *The development of ESL collocational knowledge* [Unpublished doctoral dissertation]. University of Queensland.
- Goodwin-Jones, R. (2010). Emerging technologies. From memory palaces to spacing algorithms: Approaches to second-language vocabulary learning. *Language Learning and Technology*, 14(2), 4-11.
- Goto, K. (2005). *GoTagger (Version 0.7)* [Computer software]. <http://web4u.setsunan.ac.jp/Website/GoTagger.htm#>
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology, theory, analysis and applications* (pp. 145-160). Clarendon.
- Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics* 25(1), 38-61.
- Gries, S. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics* 13(4), 403-437.
- Grucza, B., & Jaruzelska, H. (1978). Typowe błędy popelnianie przez kandydatów języku niemieckim: błędy gramatyczne i leksykalne. In F. Grucza (Ed.), *Z problematyki błędów obcojęzycznych* (pp. 124-143). WSP.
- Gyllstad, H. (2005). Words that go together well: developing test formats for measuring learner knowledge of English collocations. *International Journal of English Studies* 7(2), 127-157.
- Hadley, A. (2001). *Teaching language in context*. Heinle and Heinle.
- Halliday, M., & Sinclair, J. (1966). Beginning the study of lexis. In C. Bazell, J. Catford, M. Halliday, & R. Robins (Eds.), *In memory of J.R. Firth* (pp. 410-430). Longman.
- Hausmann, F. (1984). Wortschatzlernen ist Kollokationslernen. Zum Lshren und Lernen französischer Wortverbindungen. *Praxis des neu sprachlichen Unterrichts*, 31, 395-406.
- Heatley, A., Nation, P., & Coxhead, A. (2002). *RANGE program* [Computer software]. <http://www.victoria.ac.nz/lals/staff/paul-nation>
- Hill, J. (2000). Revising priorities: from grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 47-67). Language Teaching Publications.
- Hill, J., Lewis, M., & Lewis, M. (2000). Classroom strategies, activities, and exercises. In M. Lewis (Ed.), *Teaching Collocation: Further developments in the lexical approach* (pp. 88-116). Language Teaching Publications.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.
- Hoey, M. (1991). *Patterns of lexis in text*. Oxford University Press.
- Hopkins, G., & Bean, T. (1999). Vocabulary learning with the verbal-visual word association strategy in a native American community. *Journal of Adolescent and Adult Literacy*, 42, 274-281.
- Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Niemeyer.

- Hsu, J. (2002). *Development of collocational proficiency in a workshop on English for general business purposes for Taiwanese college students* [Unpublished doctoral dissertation]. Indiana University of Pennsylvania.
- Hsu, J., & Chiu, C. (2008). Lexical collocations and their relations to speaking proficiency of college EFL learners in Taiwan. *The Asian EFL Journal Quarterly*, 10(1), 181-204.
- Hsu, L. (2005). *The effect of lexical collocation instruction on Taiwanese college EFL learners' listening comprehension* [Unpublished master's thesis]. National Kaohsiung First University of Science and Technology.
- Huang, L. (2001). Knowledge of English collocations: An analysis of Taiwanese EFL learners. In C. Luke & B. Rubrecht (Eds.), *Texas papers in foreign language acquisition: Selected proceedings from the Texas foreign language education conference* (pp. 113-132). Texas University.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Hwang, K., & Nation, P. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System* 23(1), 35-41.
- Jaen, M. (2007). A corpus-driven design of a test for assessing the ESL collocational competence of university students. *International Journal of English Studies*, 7(2), 127-147.
- Jiang, N., & Nekrasova, T. (2007). The processing of formulaic sequences by second language speakers. *Modern Language Journal*, 91(3), 357-377.
- Jones, S., & Sinclair, J.M. (1974). English lexical collocations: a study in computational linguistics. *Catiers de Lexicologie*, 23(2), 15-61.
- Kallkvist, M. (1998). Lexical infelicity in English: The case of nouns and verbs. In K. Haastrup & A. Viberg (Eds.), *Perspectives on lexical acquisition in a second language* (pp. 149-174). Lund University Press.
- Kennedy, G. (1990). Collocations: Where grammar and vocabulary teaching meet. In S. Anivan (Ed.), *Language teaching methodology for the nineties* (pp. 215-229). Regional Language Centre.
- Keshavarz, M., & Salimi, H. (2007). Collocational competence and cloze test performance: a study of Iranian EFL learners. *International Journal of Applied Linguistics*, 17(1), 81-92.
- Kilgarriff, A., Atkins, S. & Rundell, M. (2007, July). *BNC design model past its sell-by* [Paper presentation]. University of Birmingham 4th Corpus Linguistics Conference, Birmingham, United Kingdom.
- Kjellmer, G. (1984). Some thoughts on collocational distinctiveness. In J. Aarts & W. Meijs (Eds.), *Computer corpora in English language research* (pp. 163-171). Norwegian Computing Centre for the Humanities.
- Kjellmer, G. (1987). Aspects of English collocations. In W. Meijs (Ed.), *Corpus linguistics and beyond* (pp. 133-140). Rodopi.
- Kjellmer, G. (1990). Patterns of collocability. In J. Aarts & W. Meijs (Eds.), *Theory and practice in corpus linguistics* (pp. 163-178). Rodopi.
- Kjellmer, G. (1994). *A dictionary of English collocations*. Oxford University Press.
- Korosadowicz-Struzynska, M. (1980). Word collocations in FL vocabulary instruction. *Studia Anglica Posnaniensia*, 12, 109-120.
- Koya, T. (2004). Collocation research based on corpora collected from secondary school textbooks in Japan and in the UK. *Dialogue*, 3, 7-18.

- Kuiper, K. (1996). *Smooth talkers*. Lawrence Erlbaum.
- Larsen-Freeman, D., & Long, M. (1991). *An introduction to second language acquisition research*. Longman.
- Laufer, B. (1988). *Ease and difficulty in vocabulary learning: some teaching implications* [Paper presentation]. The 22nd Annual Meeting of the International Association of Teachers of English as a Foreign Language, Edinburgh, Scotland.
- Laufer, B. (1990). Ease and difficulty in vocabulary learning: Some teaching implications. *Foreign Language Annals*, 23(2), 147-155.
- Laufer, B., & Eliasson, S. (1993). What causes avoidance in L2 learning: L1-L2 difference, L1-L2 similarity, or L2 complexity? *Studies in Second Language Acquisition*, 15(1), 35-48.
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694-716.
- Lennon, P. (1996). Getting "easy" verbs wrong at the advanced level. *International Review of Applied Linguistics*, 34, 23-36.
- Lesniewska, J., & Witalisz, E. (2007). Cross-linguistic influences on L2 and L1 collocations. *EUROSLA Yearbook*, 7, 27-48.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Language Teaching Publications.
- Lewis, M. (2000a). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 8-10). Language Teaching Publications.
- Lewis, M. (2000b). There is nothing as practical as a good theory. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 10-27). Language Teaching Publications.
- Lien, H. (2003). *The effect of collocation instruction on the reading comprehension of Taiwanese college students* [Unpublished doctoral dissertation]. Indiana University of Pennsylvania.
- Light, R. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin*, 76(5), 365-377.
- Lei, L., & Liu, D. (2018). The academic English collocation list. A corpus driven study. *International Journal of Corpus Linguistics*, 23, 216-243.
- Lin, W., Hsiao-Ching, Y., & Ho-Ping, F. (2003). English vocabulary knowledge of first-year university students: Vocabulary size and collocational knowledge. *The proceedings of the 2003 conference and workshop on FEFL and applied linguistics* (pp. 202-213). Crane Publishing Co.
- Liu, C. (1999). A study of Chinese Culture University's freshmen's collocational competence: "Knowledge" as an example. *Hwa Kang Journal of English Language and Literature*, 5, 81-99.
- Liu, L. (2002). *A corpus-based lexical semantic investigation of verb-noun miscolllocations in Taiwan learner's English* [Unpublished master's thesis]. Tamkang University.
- Liu, E., & Shaw, P. (2001). Investigating learner vocabulary: A possible approach to looking at EFL/ESL learners' qualitative knowledge of the word. *International Review of Applied Linguistics*, 39, 171-194.

- Lorenz, G. (1999). *Adjective intensification – learners versus native speakers: A corpus study of argumentative writing*. Rodopi.
- Macis, M., & Schmitt, N. (2017). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, 71, 50-59.
- Macis, M., Sonbul, S., & Alharbi, R. (2021). The effect of spacing on incidental and deliberate learning of L2 collocations. *System*, 103, 102649.
- Mackin, R. (1978). On collocations: Words shall be known by the company they keep. In P. Strevens (Ed.), *In honour of A. S. Hornby* (pp. 149-165). Oxford University Press.
- Martinez, R. (2013). A framework for the inclusion of multi-word expressions in ELT. *ELT Journal* 67(2), 184-198.
- Marton, W. (1977). Foreign vocabulary learning as problem no. 1 of language teaching at the advanced level. *Interlanguage Studies Bulletin*, 2(1), 33-57.
- Meijs, W. (1992). Inferences and lexical relations. In G. Leitner (Ed.), *New directions in English language corpora: Methodology, results, software developments* (pp. 123-152). Mouton de Gruyter.
- Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Mitchell, T. (1971). Linguistic 'goings on': Collocations and the other lexical matters arising on the syntagmatic record. *Archivum Linguisticum*, 2, 35-69.
- Monbukagakusho (2003). *The government guidelines for teaching in Japan*. www.mext.go.jp/b_m_enu/houdou/11/03/990302a/990302p.htm
- Moon, R. (1994). The analysis of fixed expressions in text. In M. Coulthard (Ed.), *Advances in written text analysis*, (pp. 117-135). Routledge.
- Moon, R. (1997). Vocabulary Connections: Multi-Words Items in English. In N. Schmitt & M. McCarthy, (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40-63). CUP.
- Morley, J. (2015). *The academic phrasebank: An academic writing resource for students and researchers*. The University of Manchester I3 Limited.
- Murphy, J. (1983, March). *Words: What goes with what?* [Paper presentation]. The 17th Annual Convention of Teachers of English to Speakers of Other Languages, Toronto, Ontario.
- Myers, J., & Chang, S. (2009). A multiple-strategy-based approach to word and collocation acquisition. *International Review of Applied Linguistics*, 47, 179-207.
- Nation, P. (2001a). How many high frequency words are there in English? In M. Gill, A.W. Johnson, L.M. Koski, R.D. Sell, & B. Wårvik (Eds.), *Language, learning and literature: Studies presented to Håkan Ringbom* (pp. 167-181). English Department Publications 4, Åbo Akademi University.
- Nation, P. (2001b). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 3-13). John Benjamins.
- Nation, P. (2008). *Teaching vocabulary: Strategies and techniques*. Heinle.
- Nation, P. & Meara, P. (2002). Vocabulary. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 35-54). Edward Arnold.

- Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle.
- Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language learning*. Oxford University Press.
- Nelson, F., & Kucera, H. (1979). *The Brown corpus: A standard corpus of present-day edited American English*. Brown University.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 233-242.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. John Benjamins.
- Noel, J. (1992). Collocation and bilingual text. In G. Leitner (Ed.), *New directions in English language corpora: Methodology, results, software developments* (pp. 345-357). Mouton de Gruyter.
- Özdem-Ertürk, Z. (2021). *Factors affecting and productive knowledge of collocations of tertiary level learners of English in Turkey* [Unpublished doctoral dissertation]. University of Hacettepe University.
- Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191-226). Longman.
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20(1), 113-138.
- Renouf, A., & Sinclair, J. (1991). Collocational frameworks in English. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics* (pp. 128-143). Longman.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on computer* (pp. 41-52). Longman.
- Robins, R. H. (1967). *A short history of linguistics*. Longman.
- Rogers, J. (2013). How many high-frequency words of English do Japanese university freshmen 'know'? *Kansai Gaikokugo University Journal of Inquiry and Research*, 97, 237-252.
- Rogers, J. (2016). *Eigo masutaa 3 zen* 英語マスター3千. Apple and Google Play App Stores.
- Rogers, J. (2017a). *What are the collocational exemplars of high-frequency English vocabulary? On identifying multi-word units most representative of high-frequency lemmatized concgrams* [Unpublished doctoral dissertation]. University of Southern Queensland.
- Rogers, J. (2017b). On the implementation of a smartphone-based English app study program. *The Proceedings for the Meijo University First International Forum*, 71-76.
- Rogers, J. (2017c). *Tōngyòng yīngyǔ dàshī* 通用英語大師. Apple and Google Play App Stores.
- Rogers, J. (2017d). *General English for Koreans*. Apple and Google Play App Stores.
- Rogers, J., Muller, A., Daulton, F., Dickinson, P., Florescu, C., Reid, G., & Stoeckel, T. (2021). The creation and application of a large-scale corpus-based academic multi-word unit list. *English for Specific Purposes*, 62(4), 142-157.

- Rodriguez, M., & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning* 50(2), 385-412.
- Rudzka, B., Channell, J., Putseys, Y., & Ostyn, P. (1981). *The word you need: Teacher's book*. Macmillan.
- Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, 18(2), 199-217.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 1-46). Cambridge University Press.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19, 17-36.
- Shin, D. (2006). *A collocation inventory for beginners* [Unpublished doctoral dissertation]. Victoria University of Wellington.
- Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 37(3), 419-441
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Sokmen, A. (1997). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy*. Cambridge University Press.
- Someya, Y. (1998). *E-lemma list*. https://lexically.net/downloads/BNC_word_lists/e_lemma.txt
- Skinner, B. (1957). *Verbal behavior*. Appleton-Century-Crofts.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative methods. *Function of Language* 2(1), 1-33.
- Taylor, C. (1983). Vocabulary for education in English. *World Language English*, 2(2), 100-104.
- Tseng, F. (2002). *A study of the effects of collocation instruction on the collocational competence of senior high school students in Taiwan* [Unpublished master's thesis]. National Taiwan Normal University.
- Twaddell, F. (1973). Vocabulary expansion in the TESOL classroom. *TESOL Quarterly*, 7(1), 61-78.
- Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences*. John Benjamins.
- University Centre for Computer Corpus Research on Language. (n.d.) *CLAWS7 tag set*. <http://ucrel.lancs.ac.uk/claws7tags.html>
- van der Meer, A. (1998). Collocations as one particular type of conventional word combinations, their definition and character. *EURALEX '98*, 1, 313-322.
- Wang, C. (2001). *A study of the English collocational competence of English majors in Taiwan* [Unpublished master's thesis]. Fu Jen Catholic University.
- Webb, S., & Nation, P. (2008). Evaluating the vocabulary load of written text. *TESOLANZ Journal*, 16, 1-10.
- Wilkins, D. (1972). *The learner's vocabulary for text analysis*. Lensing Verlag.

- Wilks, Y. (2005, June-July). REVEAL: The notion of anomalous texts in a very large corpus. *Tuscan Word Centre International Workshop*, Certosa di Pontignano, Tuscany, Italy.
- Wood, D. (2004). An empirical investigation into the facilitating role of automatized lexical phrases in second language fluency development. *Journal of Language and Learning*, 2(1), 27-50.
- Woolard, G. (2000). Collocation – encouraging learner independence. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 28-46). Language Teaching Publications.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Yamagata, S., Nakata, T., & Rogers, J. (2022). Effects of distributed practice on the acquisition of verb-noun collocations. *Studies in Second Language Acquisition*, 45(2), 291-317.
- Yorio, C. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Quarterly*, 14(4), 433-442.
- Zhang, X. (1993). *English collocations and their effect on the writing of native and non-native college freshmen* [Unpublished doctoral dissertation]. Indiana University of Pennsylvania.
- Zhou, X. (2016). A corpus-based study on high frequency verb collocations in the case of “HAVE”. *International Forum of Teaching and Studies*, 12(1), 42-50.
- Zughoul, M. (1991). Lexical choice: Towards writing problematic word lists. *International Review of Applied Linguistics*, 29, 45-58.

Index

C

chronological, 30, 34, 38, 40, 41,
45-47, 54-57, 84, 88-90, 102, 104
colligation, 18, 21-23, 33, 38, 39,
57, 59, 64, 65, 85, 95
colligational, 23, 41, 58-65, 89, 95,
103
congramming, 2-4, 20, 26, 32
congram(s), 18-20, 32, 38, 42, 46,
54, 57, 62-68, 70, 71, 73, 78-80,
83-86, 89-92
congruency, 2, 5, 10, 11, 17, 18, 26,
31, 32, 34, 37, 38, 41, 42, 45, 69-
73, 78-80, 82-86, 93, 96, 97, 102
constituent variation, 2, 18, 20
core idiom(s), 18, 42, 68, 69, 85, 92

D

dispersion, 30, 34, 38, 40, 41, 45-
49, 51-57, 84, 88, 89, 102, 104

F

figurative(s), 17, 18, 42, 68, 69, 85,
88, 92, 95

L

literal(s), 9, 17, 18, 33, 42, 68-71,
85, 88, 92, 95, 96

M

MI, 11, 29, 30, 34, 40
mutual information, 11

O

ONCE(s), 18, 42, 68

P

positional variation, 2, 18, 19

S

semantic transparency, 2, 10, 11,
18, 28, 33, 34, 38, 42, 45, 68, 70,
85, 92, 102