UNIVERSITAS
GADJAH MADA
Lexical Bundles of L1 and L2 English Writers: A Corpus-Based Study on Linguistics Research Articles
Nurinsani Yaman, Dr. Ni Gusti Ayu Roselani, M.A.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

This study examined the structures and functions of four-word lexical bundles in linguistics research articles written by L1 and L2 English writers. The data were taken from two self-compiled corpora comprising approximately 500,000 tokens each. The bundles were identified with the assistance of a corpus tool, AntConc. Frequency and dispersion threshold were applied to retrieved the list of bundles from the corpora. The bundles were then classified into the established structural and functional categories. The result revealed that L2 English writers employed more lexical bundles if compared to the L1 English writers. Structurally, the findings revealed that the majority of four-word bundles employed by L1 English writers were in PP-based and NP-based structures. In contrast, L2 English writers used more VP-based structures. However, the token frequency showed that L2 writers also used PP-based bundles almost as equal as VP-based bundles. This indicated that L2 English writers in the linguistics discipline seem to be more aware of the use of formulaic language or lexical bundles. Functionally, both groups favored text-oriented bundles. Within text-oriented category, framing signals emerge as the most dominant subcategory. The similar preference of text-oriented and framing signals between the group of writers might be related to disciplinary context of the study.

Keywords: lexical bundles, research articles, linguistics, corpus linguistics

UNIVERSITAS
GADJAH MADA
Lexical Bundles of L1 and L2 English Writers: A Corpus-Based Study on Linguistics Research Articles
Nurinsani Yaman, Dr. Ni Gusti Ayu Roselani, M.A.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Penelitian ini mengkaji struktur dan fungsi kumpulan 4-kata bundel leksikal dalam artikel penelitian linguistik yang ditulis oleh penulis L1 dan L2 bahasa Inggris. Data penelitian ini diambil dari dua korpus, yang masing-masing mencakup sekitar 500.000 token. Bundel leksikal kemudian diidentifikasi dengan bantuan AntConc, sebuah perangkat lunak korpus. Ambang batas frekuensi dan dispersi diterapkan untuk memperoleh daftar bundel leksikal dari kedua korpus tersebut. Bundel leksikal kemudian diklasifikasikan ke dalam kategori struktural dan fungsional yang telah ditetapkan. Hasil penelitian menunjukkan bahwa penulis L2 bahasa Inggris menggunakan lebih banyak bundel leksikal dibandingkan dengan penulis L1. Secara struktur, sebagian besar bundel leksikal yang digunakan oleh penulis L1 berbasis frasa preposisi (*PP-based*) dan frasa nomina (*NP-based*). Sebaliknya, penulis L2 lebih banyak menggunakan bundel berbasis frasa verba (*VP-based*). Namun, frekuensi token struktur bundel leksikal menunjukkan bahwa penulis L2 juga menggunakan bundel berbasis frasa preposisi (*PP-based*) yang hampir sama seringnya dengan frasa verba (*VP-based*). Hal ini mengindikasikan bahwa penulis L2 bahasa Inggris di bidang linguistik memiliki kedasaran dalam penggunaan bahasa formulaik atau bundel leksikal. Secara fungsi, kedua kelompok penulis lebih banyak menggunakan bundel *text-oriented*. Dalam kategori ini, *framing signals* merupakan subkategori yang paling dominan. Kesamaan preferensi terhadap kategori dan subkategori pada fungsi bundel leksikal di antara kedua kelompok penulis tersebut mungkin berkaitan dengan konteks disiplin ilmu dalam penelitian ini.

Kata Kunci: bundel leksikal, artikel penelitian, linguistik, linguistik korpus