



Linguistik, Terjemahan, Sastra (LINGTERSA)

Journal homepage: <https://talenta.usu.ac.id/lingtersa>



Analisis Frekuensi Kata dalam Cerpen: Studi Linguistik Korpus dalam Karya Sastra

Maytha Carissa Ariyanti^{*1}, Pratomo Widodo², Pangesti Wiedarti³, Titik Sudartinah⁴

^{1,2,3,4}Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

*Corresponding Author: maythacarissa@gmail.com

ARTICLE INFO

Article history:

Received 22 Dec 2024

Revised 02 Jan 2025

Accepted 03 Feb 2025

Available online

<https://talenta.usu.ac.id/lingtersa/index>

E-ISSN: 2964-1713

P-ISSN: 2775-5622

ABSTRACT

*This study aims to analyze the frequency of words in the translated short story *The Tale of Johnny Town-Mouse* by Beatrix Potter translated by Arya Bisma into *The Tale of Johnny Town-Mouse*. A qualitative approach is used to identify patterns of word usage as a reflection of the themes, characters, and narrative elements of the story. Then the data analysis technique was carried out using the web-based software Voyant Tools, which systematically counts the number of occurrences of each word in the text. The results showed that words with the highest frequency such as "Timmy," "Willie," and "Johnny" reflected the story's focus on the interactions between the main characters. The word "keranjang" indicates the importance of physical elements associated with rural life, while "town" and "mouse" emphasize the setting and theme of city life. In addition, descriptive words such as "suara" and "kembali" reinforce the atmosphere and plot of the story. This study shows that corpus analysis can reveal linguistic patterns that reflect important themes and elements in literary works. This approach not only helps to understand language structure and style, but also provides insight into the social and cultural context that influenced the story. Thus, this study emphasizes the relevance of the corpus linguistics approach in literary analysis*

Keywords: *corpus analysis, word frequency, Voyant Tools, short stories*

ABSTRAK

Penelitian ini bertujuan untuk menganalisis frekuensi kata dalam cerita pendek terjemahan *The Tale of Johnny Town-Mouse* karya Beatrix Potter yang diterjemahkan oleh Arya Bisma menjadi *Kisah Johnny Town-Mouse*. Pendekatan kualitatif digunakan untuk mengidentifikasi pola penggunaan kata sebagai cerminan tema, karakter, dan elemen naratif cerita. Kemudian teknik analisis data dilakukan dengan menggunakan perangkat lunak berbasis web Voyant Tools, yang secara sistematis menghitung jumlah kemunculan setiap kata dalam teks. Hasil penelitian menunjukkan bahwa kata-kata dengan frekuensi tertinggi seperti "Timmy," "Willie," dan "Johnny" mencerminkan fokus cerita pada interaksi antar karakter utama. Kata "keranjang" mengindikasikan pentingnya elemen fisik yang terkait dengan kehidupan pedesaan, sementara "town" dan "mouse" menekankan latar cerita dan tema kehidupan kota. Selain itu, kata-kata deskriptif seperti "suara" dan "kembali" memperkuat suasana serta alur cerita. Penelitian ini menunjukkan bahwa analisis korpus dapat mengungkap pola linguistik yang mencerminkan tema dan elemen penting dalam karya sastra. Pendekatan ini tidak hanya membantu memahami struktur dan gaya bahasa, tetapi juga memberikan wawasan tentang konteks sosial dan budaya yang memengaruhi cerita. Dengan demikian, studi ini mempertegas relevansi pendekatan linguistik korpus dalam analisis sastra.

Kata Kunci: analisis korpus, frekuensi kata, Voyant Tools, cerita pendek



**This work is licensed
under a Creative
Commons Attribution-
NonCommercial-
NoDerivatives 4.0
International.**

<http://doi.org/10.26594/re-gister.v6i1.idarticle>

1. Pendahuluan

Bahasa mencerminkan pola komunikasi yang digunakan dalam masyarakat. Menurut Sapir (1921), bahasa adalah metode utama yang digunakan manusia untuk mengomunikasikan ide, emosi, dan keinginan melalui sistem simbol yang diproduksi secara sukarela. Sebagai sistem simbolik, bahasa tidak hanya berfungsi sebagai alat komunikasi tetapi juga sebagai media untuk mengekspresikan budaya, nilai, dan perspektif masyarakat. Dalam konteks karya sastra, khususnya cerpen, bahasa menjadi media utama untuk menyampaikan pesan, suasana, dan karakterisasi yang memikat. Pilihan kata, diksi, dan istilah yang digunakan oleh penulis tidak hanya menggambarkan isi cerita, tetapi juga membangun makna yang beresonansi dengan pembaca. Penggunaan kata-kata tertentu yang sering muncul dalam sebuah cerpen dapat mencerminkan tema utama, konflik, atau gagasan yang dianggap relevan dan menarik. Selain itu, istilah-istilah tersebut sering kali mencerminkan pola pikir atau isu sosial yang sedang dihadapi masyarakat pada waktu tertentu. Dengan demikian, analisis bahasa dalam cerpen tidak hanya membantu memahami isi cerita, tetapi juga memberikan wawasan tentang realitas sosial dan budaya yang memengaruhi pembentukan cerita tersebut.

Dalam era digital saat ini, analisis bahasa dalam karya sastra dapat dilakukan dengan bantuan teknologi seperti analisis korpus. Menurut Hizbullah, Fazlurrahman, dan Fauziah (2016) Korpus adalah kumpulan data, baik dalam bentuk teks cetak maupun digital, yang memuat berbagai informasi kebahasaan, seperti kosakata, struktur kalimat, makna, hingga wacana, dan dapat digunakan untuk tujuan penelitian. Kemudian, menurut O’Keeffe et al. (2007) *A corpus is a collection of texts, written or spoken, which is stored on a computer* ‘Korpus adalah kumpulan teks, tertulis atau lisan, yang disimpan di komputer’. Pendekatan ini memungkinkan peneliti untuk mengidentifikasi pola penggunaan kata-kata dan istilah dalam cerpen secara objektif dan sistematis. Dengan menggunakan kumpulan teks digital dari berbagai cerpen, peneliti dapat mengeksplorasi kecenderungan linguistik yang berkontribusi pada daya tarik cerita di mata pembaca. Studi semacam ini tidak hanya bermanfaat untuk memahami gaya penulisan, tetapi juga untuk mengungkap hubungan antara bahasa, budaya, dan minat pembaca dalam dunia sastra.

Penelitian ini meneliti tentang bagaimana tren linguistik yang muncul dapat dipahami melalui analisis kata-kata dengan frekuensi tinggi dalam cerita pendek hasil terjemahan, sehingga tujuan penelitian ini adalah untuk menganalisis kata berdasarkan tingkat kemunculannya (frekuensi penggunaan) yang tinggi dalam cerita pendek hasil terjemahan untuk memahami tren linguistik yang muncul. Sinclair dalam Harahap, Sinar, dan Prihantoro (2024) menyatakan bahwa salah satu inovasi terkini dalam studi linguistik adalah pemanfaatan korpus, yaitu himpunan teks digital yang disusun berdasarkan kriteria tertentu untuk merepresentasikan suatu bahasa atau variasi bahasa tertentu. Kemunculan cabang ilmu linguistik korpus, yang dianggap relatif "baru," turut meramaikan perkembangan linguistik modern. Ilmu ini secara khusus mengkaji bahasa menggunakan kumpulan data yang alami dan nyata sesuai konteks penggunaannya, baik berupa teks tertulis maupun transkripsi dari ujaran lisan (Adolphs, 2006: 137).

Kajian tentang korpus linguistik merupakan topik yang menarik untuk dieksplorasi dan telah menjadi fokus penelitian oleh beberapa peneliti sebelumnya. Namun, pada penelitian-penelitian sebelumnya membahas tentang penggunaan korpora pada studi sinonim, seperti penelitian yang telah dilakukan oleh Aroonmanakun (2015) dengan judul *“Quick or Fast: A Corpus Based Study of English Synonyms”*. Selanjutnya, penelitian

dengan kajian neologisme pendekatan linguistik korpus, contoh penelitian ini adalah penelitian yang dilakukan oleh Wahyuningsih (2022) dengan judul “*Istilah Baru Era Pandemi COVID-19 di Bahasa Indonesia dan Bahasa Korea: Kajian Neologisme Pendekatan Linguistik Korpus*” serta penelitian mengenai pemanfaatan linguistik korpus dalam menentukan kata berfrekuensi tinggi pada sebuah buku. Penelitian ini telah dilakukan oleh Wahyuningtyas dan Kesuma (2021) dengan judul “*Pemanfaatan Linguistik Korpus dalam Menentukan Kata Berfrekuensi Tinggi pada Buku Sahabatku Indonesia BIPA 1*”. Penelitian ini berbeda dengan penelitian-penelitian sebelumnya dalam hal objek dan subjek penelitian, serta alat analisis yang digunakan. Penelitian ini berfokus pada analisis teks cerita pendek dengan mengidentifikasi kata-kata yang paling sering muncul menggunakan perangkat lunak berbasis web, Voyant Tools. Voyant Tools merupakan aplikasi yang dirancang untuk menganalisis teks dalam jumlah besar.

Bahasa dalam cerpen tidak hanya berfungsi sebagai media komunikasi, tetapi juga mencerminkan budaya, nilai, dan isu sosial masyarakat. Analisis frekuensi kata pada cerpen hasil terjemahan penting dilakukan untuk memahami tren linguistik, tema, dan gagasan yang relevan, sekaligus mengungkap pengaruh lintas budaya dalam teks. Dengan kemajuan teknologi, alat seperti Voyant Tools memungkinkan analisis teks secara objektif dan sistematis dalam skala besar, mengatasi keterbatasan metode manual. Penelitian ini menjadi penting karena mengisi kesenjangan studi linguistik korpus pada cerpen hasil terjemahan, memberikan wawasan baru tentang hubungan bahasa, budaya, dan pembaca, serta memperkaya metode analisis berbasis teknologi dalam kajian sastra.

2. Metode Penelitian

Pendekatan Penelitian

Penelitian ini dilakukan dengan pendekatan kualitatif, yang bertujuan untuk memahami fenomena secara mendalam berdasarkan data yang dikumpulkan. Menurut Moleong (2004), penelitian kualitatif bertujuan untuk memperoleh pemahaman yang lebih mendalam mengenai fenomena yang dialami oleh peneliti, mencakup aspek-aspek seperti perilaku, motivasi, dan tindakan. Data yang dianalisis dalam penelitian ini adalah kata-kata yang memiliki tingkat kemunculan tinggi (frekuensi kata).

Sumber Data

Sumber data utama dalam penelitian ini adalah sebuah cerpen terjemahan karya Beatrix Potter berjudul *The Tale of Johnny Town-Mouse*. Cerpen ini telah diterjemahkan ke dalam bahasa Indonesia oleh Arya Bisma dengan judul “Kisah Johnny Town-Mouse”.

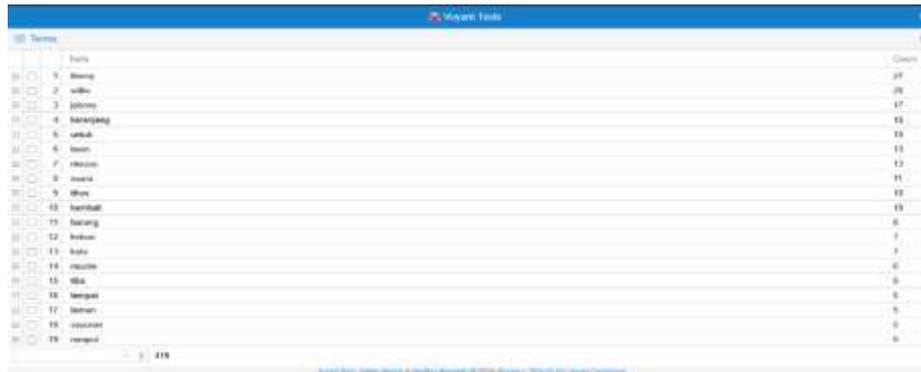
Teknik Pengumpulan Data

Tahapan pengumpulan data dilakukan sebagai berikut:

1. Penyalinan Teks: Teks cerpen terjemahan disalin secara manual dari sumbernya.
2. Pembersihan Data: Teks yang telah disalin kemudian dibersihkan dari elemen non-teks seperti: iklan, metadata, elemen-elemen lain yang tidak relevan dengan isi teks utama.

Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan menggunakan perangkat lunak berbasis web bernama Voyant Tools. Langkah-langkah analisis data secara rinci adalah sebagai berikut:

Gambar 2. Visualisasi kata-kata yang muncul dengan frekuensi tinggi dalam bentuk *terms*

Pembahasan

Hasil analisis menggunakan Voyant Tools menunjukkan bahwa kata-kata dengan frekuensi tertinggi adalah "Timmy" (27 kali), "Willie" (26 kali), dan "Johnny" (17 kali). Ketiga kata ini merujuk pada nama karakter utama dalam cerita. Frekuensi yang tinggi menunjukkan bahwa cerita sangat berpusat pada interaksi antara karakter-karakter ini. Hal ini mencerminkan fokus narasi pada hubungan, konflik, dan dinamika yang terjadi di antara mereka. Sebagai contoh, "Timmy" dan "Willie" kemungkinan besar mewakili karakter utama yang memiliki pandangan atau pengalaman berbeda mengenai kehidupan di kota dan pedesaan, yang menjadi tema utama cerita.

Kata "keranjang" yang muncul sebanyak 16 kali memberikan petunjuk tentang elemen fisik atau simbolis yang signifikan dalam cerita. Dalam konteks cerpen, "keranjang" mungkin merepresentasikan alat atau simbol dari kehidupan pedesaan yang erat dengan tokoh Timmy. Kemunculan kata ini yang cukup sering menunjukkan pentingnya objek tersebut dalam alur cerita, baik sebagai latar, alat untuk membangun konflik, atau sebagai metafora kehidupan sederhana.

Selanjutnya, kata "untuk" yang muncul sebanyak 15 kali adalah bagian dari struktur bahasa yang mendukung hubungan antarfrasa dan kalimat. Meskipun terlihat sebagai elemen kebahasaan yang umum, frekuensi kemunculannya tetap penting karena menunjukkan bagaimana hubungan antara tindakan dan tujuan disampaikan dalam cerita.

Kata "town" dan "mouse," masing-masing muncul 13 kali. Hal ini mencerminkan tema utama cerita yang berkaitan dengan kehidupan tikus di kota. Kata-kata ini memperkuat latar cerita dan menjadi penanda penting untuk menggambarkan lingkungan urban dan interaksi para tokoh di sana. Sementara itu, kata "suara" (11 kali) dan "tikus" (10 kali) menekankan deskripsi suasana yang memperkaya narasi. "Suara" berfungsi untuk membangun atmosfer cerita, sedangkan "tikus" mengacu pada karakter utama cerita.

Kata "kembali" yang juga muncul sebanyak 10 kali menunjukkan adanya alur cerita yang melibatkan pergerakan atau perpindahan antara tempat atau situasi tertentu. Hal ini dapat mengindikasikan konflik dan penyelesaiannya yang melibatkan perjalanan atau perubahan dari satu tempat ke tempat lain.

Kata-kata lain seperti "burung," "kebun," dan "rumput," yang muncul masing-masing beberapa kali, menambahkan detail latar pedesaan, memperkuat kontradiksi antara kehidupan kota dan desa yang menjadi tema cerita. Sedangkan kata "musim," "tiba," dan "tempat" memberikan konteks waktu dan ruang yang lebih spesifik untuk mendukung narasi.

Kesimpulan

Berdasarkan hasil analisis menggunakan Voyant Tools, ditemukan bahwa kata-kata dengan frekuensi tertinggi, seperti "Timmy," "Willie," dan "Johnny," merujuk pada karakter utama, menegaskan bahwa cerita berpusat pada interaksi dan dinamika di antara mereka. Selain itu, kemunculan kata-kata seperti "keranjang," "town," dan "mouse" menggambarkan elemen-elemen narasi utama, yaitu tema kehidupan di kota dan desa, serta perbedaan pengalaman di kedua lingkungan tersebut. Kata-kata lain seperti "suara," "tikus," dan "kembali" menambahkan kedalaman pada latar cerita dan suasana, memperkuat konflik dan penyelesaian cerita.

Hasil penelitian ini menunjukkan bahwa analisis kata dengan frekuensi tinggi dapat mengungkap tema utama, konflik, dan elemen penting dalam cerita, sekaligus memberikan wawasan tentang hubungan bahasa dengan budaya, nilai, dan realitas sosial. Dengan demikian, pendekatan berbasis korpus ini efektif untuk memahami karakteristik linguistik dan daya tarik sebuah karya sastra.

Daftar Pustaka

- Aroonmanakun, V. (2015). Quick or Fast: A Corpus Based Study of English Synonyms. *Language Education and Acquisition Research Network (LEARN) Journal*. 8(1)
- Harahap, F.S., Sinar, T.S., & Prihantoro. (2024). Analisis Verba ‘Ajar’ dalam Korpus Bahasa Indonesia: Studi Linguistik Korpus. *Lingtersa: Linguistik, Terjemahan, Sastra*
- Hizbullah, N., Fazlurrahman, & Fauziah, F. (2016). Linguistik Korpus dalam Kajian dan Pembelajaran Bahasa Arab di Indonesia. *Prosiding Konferensi Nasional Bahasa Arab: KONASBARA*.
- Johns, J. L., & Wilke, K. H. (2018). *High Frequency Words: Some Ways to Teach and Help Students Practice and Learn Them*. *Texas Journal of Literacy Education*. 6(1)
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press
- Oktriani, A.D., Maisarah, I., & Erniyati, Y. (2022). Translation of the English nominal suffixes into Indonesian in Louisa May Alcott’s *Little Woman: The equivalent and shift category*. *JASL: Journal of Applied Studies in Language*. 6(1)
- Siagian, E. N. (2020). Kata Berfrekuensi Tinggi dalam Pembelajaran BIPA Pemula. *Ranah: Jurnal Kajian Bahasa*, 9(2), 188–201
- Sinclair, J. (Ed.) (1987). *Looking up. An account of the COBUILD Project in lexical computing*. London: Collins Corbulid
- Wahyuningsih, S. (2022). Istilah Baru Era Pandemi COVID-19 di Bahasa Indonesia dan Bahasa Korea: Kajian Neologisme Pendekatan Linguistik Korpus. *JLA: Jurnal Lingua Applicata*. 5(2)
- Wahyuningtyas, D., & Kesuma, T.M.J. (2021). Pemanfaatan Linguistik Korpus dalam Menentukan Kata Berfrekuensi Tinggi pada Buku Sahabatku Indonesia BIPA 1. *JBIPA: Jurnal Bahasa Indonesia bagi Penutur Asing (JBIPA)*. 3(2)