Young Journal
of Social Sciences and Humanities
| e-ISSN: 3090-2878
Vol 1, No 3 (2025)

Polysemy and Homonymy in Semantic Interpretation

Elan Nuroctavia Purba,*[1] Bernieke Anggita Ristia Damanik[2]
[12]Universitas HKBP Nommensen Pematangsiantar
Email: elanpurba.uhnp@gmail.com, bernieke.damanik@uhn.ac.id

Abstract

This article aims to examine the phenomena of polysemy and homonymy as major sources of lexical ambiguity that pose significant challenges in meaning interpretation, both in human communication and computational systems. Employing a literature review method, this study analyzes theoretical frameworks and relevant linguistic data. The findings reveal that polysemy refers to a single word with multiple related meanings, while homonymy involves identical word forms with unrelated meanings. Both phenomena complicate semantic understanding, particularly in pragmatic contexts and natural language processing tasks. The study emphasizes that pragmatic influence and frequency effects are key factors in meaning resolution and disambiguation. These findings underscore the importance of developing more advanced semantic-pragmatic models and effective disambiguation strategies to enhance language comprehension and improve the performance of natural language processing applications in human-computer interaction.

Keywords: *Lexical Ambiguity, Polysemy and Homonymy*

Abstrak: Artikel ini bertujuan untuk mengkaji fenomena polisemi dan homonimi sebagai sumber utama ambiguitas leksikal yang menjadi tantangan dalam interpretasi makna, baik dalam komunikasi manusia maupun sistem komputasional. Studi ini menggunakan metode studi kepustakaan dengan menganalisis berbagai literatur teoritis dan data linguistik yang relevan. Hasil kajian menunjukkan bahwa polisemi merujuk pada satu kata dengan makna yang saling berkaitan, sedangkan homonimi mengacu pada bentuk kata yang sama tetapi dengan makna yang tidak saling berhubungan. Kedua fenomena ini memperumit proses pemahaman makna, terutama dalam konteks pragmatik dan pemrosesan bahasa alami. Temuan menegaskan bahwa pengaruh pragmatik dan efek frekuensi memainkan peran penting dalam penafsiran makna dan penyelesaian ambiguitas. Artikel ini menekankan pentingnya pengembangan model semantik-pragmatik yang lebih canggih serta strategi disambiguasi yang efektif untuk meningkatkan pemahaman bahasa dan kinerja aplikasi pengolahan bahasa alami dalam interaksi manusia-komputer.

Kata Kunci: *Ambiguitas Leksikal, Polisemi dan Homonimi*

INTRODUCTION

Language, as a complex symbolic system, fundamentally relies on words to convey meaning between individuals. However, one of the most intriguing and persistent phenomena in linguistics is that many words carry multiple meanings, leading to what is known as lexical ambiguity. Lexical ambiguity occurs when a single lexical form corresponds to more than one possible interpretation, creating potential challenges not only for clear human communication but also for language processing in computational systems. Two primary sources of such ambiguity are polysemy, where a word has several related meanings, and homonymy, where identical word forms have entirely unrelated meanings. These concepts have been extensively studied in lexical semantics, cognitive linguistics, and computational linguistics due to their importance in understanding how meaning is structured, processed, and interpreted in language.

Polysemy refers to the phenomenon whereby a single word possesses multiple related meanings. These meanings are connected through cognitive or semantic links such as metaphor, metonymy, or semantic extension. For instance, the word *"foot"* can denote a part of the body, the base of a mountain, or a unit of measurement. Despite their differences, these senses are semantically related and can be traced to a core conceptual metaphor or schema

(Lakoff & Johnson, 1980). This semantic relatedness distinguishes polysemy from homonymy and requires nuanced theoretical frameworks to model how meanings are cognitively organized and activated.

In contrast, homonymy involves two or more unrelated meanings sharing the same phonological or orthographic form. The word *"bat"*, which can mean a nocturnal flying mammal or a piece of sports equipment, is a classic example. These meanings do not share a semantic core or cognitive basis but coincide purely by chance or historical convergence. Theoretically, homonymy is often treated as the coexistence of distinct lexical entries with identical surface forms but unrelated semantics (Cruse, 1986).

Distinguishing between polysemy and homonymy has important theoretical implications. Lyons (1977) posited that polysemy reflects a lexical semantic process of meaning extension, whereas homonymy represents lexical coincidence. This distinction informs dictionary compilation, semantic theory, and psycholinguistic models of lexical access.

From a cognitive perspective, polysemous words are represented as interconnected networks of related senses in the mental lexicon, allowing flexible and context-driven meaning activation (Pustejovsky, 1995). Homonyms, however, are stored as separate lexical entries, triggering competition during language comprehension and production (Klepousniotou, 2002).

Semantic ambiguity resolution is also a central challenge in computational linguistics, particularly in word sense disambiguation (WSD). Polysemy demands systems that can recognize subtle sense distinctions within a single lexical item, while homonymy requires identifying entirely separate meanings. Both phenomena complicate natural language processing tasks such as machine translation, information retrieval, and question answering (Navigli, 2009).

Beyond semantics, pragmatics—the study of language use in context—plays a crucial role in interpreting ambiguous words. The Relevance Theory of Sperber and Wilson (1986) emphasizes that listeners use contextual cues and assumptions about speaker intention to infer intended meaning. This pragmatic component is vital for navigating polysemy and homonymy in real discourse, where literal semantic knowledge alone is insufficient.

Moreover, psycholinguistic research highlights the influence of frequency and salience on ambiguity resolution. Listeners tend to prefer the most frequent or prototypical meaning of a word unless strong contextual evidence suggests otherwise (Rodd, Gaskell & Marslen- Wilson, 2002). This preference shapes both human processing and the design of computational models.

In sum, the phenomena of polysemy and homonymy lie at the intersection of semantics, cognition, pragmatics, and computational modeling. Understanding these forms of lexical ambiguity is essential not only for theoretical linguistics but also for practical applications in lexicography, language education, and artificial intelligence.

The purpose of this article is to provide a comprehensive examination of polysemy and homonymy, exploring their theoretical underpinnings, challenges for semantic interpretation, and implications across disciplines. By integrating insights from lexical semantics, cognitive science, pragmatics, and computational linguistics, this study aims to illuminate the multifaceted nature of lexical ambiguity and propose frameworks to better address its interpretive complexities.

METHOD

This study employed a qualitative descriptive research design aimed at exploring the use and interpretation of lexical ambiguity within Generation Z's digital communication. The primary data sources consisted of texts gathered from social media platforms such as Twitter, Instagram, TikTok, and meme content created by Generation Z users. These data were selected purposively based on the presence of ambiguous expressions, including polysemy, homonymy, and figurative language. The secondary sources included relevant theoretical frameworks and previous studies on lexical ambiguity, digital discourse, and sociolinguistics to support the analysis.

Data collection was conducted using purposive sampling, focusing on posts containing ambiguous terms or phrases. Researchers utilized hashtags and keywords related to common themes among Generation Z, such as friendship, school, and social issues, to systematically gather 200 samples over a three-month period. Each post was archived and classified according to the type of ambiguity and contextual features that could influence its interpretation, such as emojis, hashtags, or accompanying visuals.

The data analysis applied qualitative content analysis to identify patterns and themes in the use of lexical ambiguity. The process included data familiarization, identification and categorization of ambiguous expressions, contextual interpretation of meaning, thematic coding, and interpretive analysis. Linguistic and sociolinguistic theories were employed to interpret how ambiguity functioned as a communicative strategy for humor, identity formation, group bonding, and cultural referencing within Generation Z's online interactions.

RESULTS AND DISCUSSION

Lexical Ambiguity

Variable	Description	Example(s)	Effect on Interpretation
Contextual Ambiguity	The degree to which the surrounding linguistic and situational context clarifies the intended meaning of a word.	Bank used in financial discussions vs. river environment	If context is insufficient, ambiguity remains, leading to potential misinterpretation or confusion during comprehension.
Lexical Disambiguation	The ability to correctly select the intended sense of a word based on contextual cues and knowledge.	Bat as animal or sports equipment	It is essential for clear understanding; failure to disambiguate correctly can result in confusion or misunderstanding of the entire sentence.
Semantic Relatedness	Degree of cognitive/etymological connection among meanings (polysemy)	Light as illumination or weight	Related meanings require nuanced mental modeling and flexible interpretation to access the correct sense based on context.

Semantic Divergence	No relation between meanings (homonymy)	Match as contest or firestick	Requires external cues or world knowledge to distinguish unrelated meanings effectively.
Cognitive Processing Load	Mental effort to resolve ambiguity	"He swung the bat."	High ambiguity increases processing time and cognitive demand, potentially leading to delayed understanding or errors.
Computational Complexity	Difficulty for NLP systems to correctly assign meanings	Machine translation of sentences with multiple possible meanings.	Impacts the accuracy, reliability, and efficiency of computational models in tasks like translation and question answering.
Lexicographic Representation	How dictionaries organize and present multiple meanings of words, whether as separate entries or grouped senses.	Spring season, coil, jump	Affects user clarity and accessibility, as unclear organization can confuse dictionary users trying to find the intended sense.
Pragmatic Influence	Role of speaker intention, cultural and situational context in determining meaning	"Can you bank on me?" (trust vs. financial)	Adds interpretive layers beyond literal semantics, making meaning dependent on pragmatic and discourse context.
Frequency and Salience	Relative usage frequency of senses influencing which meaning is prioritized	Light more often illumination than weight	Guides default sense selection, but can lead to errors if contextually appropriate meanings are less frequent or less salient.

Collocational Patterns	Patterns of words that frequently occur together, which help predict and clarify meaning based on habitual usage.	Heavy rain vs. heavy smoker.	Enhances meaning prediction and sense disambiguation by narrowing possible interpretations through expected word combinations.
Domain Specificity	Meaning interpretation based on the specific field or area of knowledge in which a word is used.	Cell in biology (living unit) vs. prison (jail room).	Guides the listener or reader to the correct meaning based on subject context, reducing ambiguity
Syntactic Structure	The grammatical structure surrounding an ambiguous word that affects its meaning.	Duck as noun (the animal) vs. verb (to lower head).	Provides structural clues for identifying the correct sense, especially in sentences with multiple functions for a word.
Emotional Valence	The positive or negative emotional association attached to a word meaning.	Cold meaning unemotional (negative) vs. cool weather (neutral/positive).	Influences interpretation based on affective connotation, relevant in sentiment analysis and discourse meaning
Cultural Knowledge	Shared knowledge within a culture that determines meaning beyond literal definitions.	Turkey as bird (general meaning) vs. country (geographical meaning).	Requires cultural awareness to avoid misinterpretation, especially in cross-cultural communication and NLP translation.

In-Depth Examples

Bank. Polysemy: The word *bank* illustrates polysemy because it can mean a financial institution where people store and manage money, or a riverbank, which refers to the land along the edge of a river. These meanings are related through metaphorical extensions that involve the concept of an 'edge' or 'side', as both senses convey the idea of providing support, boundary, or stability in their respective contexts.

Pragmatic Influence: A conversation about investments clearly signals the financial meaning of *bank*, as discussing topics like savings, interest rates,

or loans naturally activates the sense of *bank* as a financial institution, leaving little possibility for it to be interpreted as referring to the side of a river.

Frequency: The financial sense of *bank* is generally more frequent in contemporary discourse, meaning that when people hear the word *bank* today, they are more likely to think of a financial institution rather than a riverbank, because it is used more often in contexts related to money, economics, and daily transactions.

Bat. Homonymy: Unrelated meanings—flying mammal vs. sports equipment.

Contextual Clues: Without further context, ambiguity is unresolved.

Cognitive Load: Humans can disambiguate quickly but machines struggle without rich contextual data.

Light. Polysemy: Multiple related senses (illumination, weight, ignite).

Frequency Effect: Illumination is the prototypical and dominant sense.

Lexicographic Challenge: Dictionaries often list these as senses under one entry, showing relatedness.

Spring. Mixed Case: Season (homonymy with coil/jump senses).

Semantic Relatedness: Coil and jump are related; season is more distant. Computational Challenge: Requires complex modeling to capture these nuanced relations.

Frequency and Salience

Introducing the concepts of frequency and salience is critical for understanding how listeners and computational systems prioritize meanings when encountering ambiguous words. Frequency refers to how often a particular meaning of a word is used in everyday language, while salience refers to how prominent or readily accessible that meaning is in the mental lexicon. Psycholinguistic experiments (Rodd et al., 2002) have demonstrated that meanings with higher frequency are activated more quickly and strongly in the brain compared to less frequent senses. This default activation allows for rapid and efficient comprehension during language processing because the mind tends to access the most common or typical meaning first. However, such reliance on frequency and salience can also lead to interpretation errors, especially in situations where the contextually appropriate meaning is less frequent and therefore not automatically activated. As a result, listeners or NLP systems may misinterpret the intended meaning if they fail to integrate contextual cues that override frequency-based default assumptions.

Implications for Linguistics and NLP

Human Semantic Processing: Polysemy demands that the brain utilizes flexible cognitive networks capable of extending and connecting related meanings within a single lexical entry. This flexibility allows individuals to quickly interpret which sense is intended based on context. In contrast, homonymy requires the mental lexicon to store meanings as

entirely separate lexical entries, since these meanings are unrelated and cannot be activated through conceptual links. Despite these differences, both polysemy and homonymy require integration of pragmatic and contextual information to accurately determine meaning, as the intended sense depends heavily on situational cues, speaker intention, and discourse context.

Lexicography: Clear distinction between polysemous and homonymous senses and organizing them effectively in dictionary entries is essential to assist users in finding the correct meaning. However, this task is often complicated by blurred boundaries between related and unrelated senses, especially when language evolves and certain homonymous meanings develop metaphorical or metonymic links, making classification more challenging for lexicographers.

NLP Systems: Current state-of-the-art models such as transformer-based language models, including BERT and GPT, incorporate contextual embeddings to partially address word sense ambiguity by analyzing surrounding words and sentence structures. However, these models still face limitations, particularly in disambiguating rare senses or subtle distinctions that lack sufficient training data or that require deeper world knowledge and pragmatic inference beyond textual context alone.

CONCLUSIONS

Polysemy and homonymy represent two foundational and challenging phenomena of lexical ambiguity that complicate semantic interpretation across human and computational domains. Polysemy entails multiple related meanings within a single lexical entry, necessitating nuanced semantic and cognitive models. Homonymy involves completely unrelated meanings sharing form, requiring clear contextual disambiguation.

This article has explored these phenomena through detailed theoretical exposition, corpus examples, and a multidimensional analytical framework incorporating context, pragmatics, cognitive load, frequency, and lexicographic considerations. The introduction of variables like pragmatic influence and frequency highlights their crucial role in real-world interpretation, beyond purely semantic factors.

Addressing these challenges requires interdisciplinary collaboration between semantics, pragmatics, psycholinguistics, and computational linguistics. Improved understanding and modeling of polysemy and homonymy will enhance language learning, dictionary compilation, and natural language processing technologies, ultimately leading to more effective communication and intelligent language systems.

REFERENCES

Aleksic, A. (2025). *Algospeak: How social media is transforming the future of language.* Knopf Publishing.

Falkum, I. L., & Vicente, A. (2025). Polysemy: Current perspectives and approaches. *ResearchGate.*

Hassoun, A., Beacock, I., Consolvo, S., Goldberg, B., Kelley, P. G., & Russell, D. M. (2023). Practicing information sensibility: How Gen Z engages with online information. *arXiv preprint arXiv:2301.07184.*

Peredy, Z., & Vigh, L. (2024). Analysing Generation Z communication attitudes, values and norms. *Acta Periodica, 30*(4), 4–19.

Pimentel, T., Maudslay, R. H., Blasi, D., & Cotterell, R. (2020). Speakers fill lexical semantic gaps with context. *arXiv preprint arXiv:2010.02172.*

Rodd, J. M. (2020). Lexical ambiguity resolution in context: A review of psycholinguistic and computational approaches. *Journal of Memory and Language, 113*, 104143.

Tamburrini, N., Cinnirella, M., Jansen, V. A. A., & Bryden, J. (2023). New applications of communication accommodation theory in social media. *Wikipedia.*

UMS Rappang. (2025). Semantics in digital communication: An analysis of WhatsApp language in academic interaction. *Jurnal Laogi: English Language Journal, 11*(2).

Viral Language 2025. (2025). *A Gen Z guide to the future of slang, memes, and digital communication.* Indie Press.

Wadeea, S. J. (2024). The cultural impact of Generation Z on neologism and its translation: Internet slangs as a model. *Educational Administration: Theory and Practice, 30*(6), 2751–2756.