Comparing Lexical Structures: A Network Analysis of WordNet and LLM-Generated Synonyms

Sanzhar Sailaubek
sanzhar.sailaubek@studenti.unitn.it

June 5, 2025

Abstract

Understanding lexical semantics is fundamental to natural language processing, and synonym networks offer a graph-based approach to capture semantic relationships between words. This study investigates and compares synonym networks generated by large language models (LLMs) with those derived from WordNet, a widely used human-curated lexical database. Results highlight LLMs' suitability for flexible, context-driven tasks and WordNet's strength in controlled, accurate synonymy.

1 Introduction

Synonym networks model lexical relationships and semantic similarity across natural language. WordNet, a widely used expert-curated lexical database, has served as a key resource in computational linguistics. These networks often exhibit scale-free properties and semantic clustering, as shown by (Steyvers & Tenenbaum, 2005), revealing the latent organization of language.

The emergence of large language models (LLMs) like Mistral and GPT has enabled dynamic generation of lexical associations based on contextual learning. LLMs have been shown to encode both factual and semantic knowledge (Petroni et al., 2019), with embedding-based architectures effectively capturing word meaning (Mikolov, Chen, Corrado, & Dean, 2013). Our goal is to understand how curated and AI-generated systems differ in representing synonymy.

2 Experimental Design

2.1 Data Collection

We selected 240 high-frequency English adjectives. For the LLM-based network, we used Mistral-small, prompting it to generate 3–10 synonyms per word. This produced more 938 synonym terms, later expanded with further synonym queries. For WordNet, we used the NLTK interface to retrieve up to 10 synonyms per adjective, focused on adjectival senses.

In both networks, undirected edges were created between base words and their synonyms.

3 Comparative Analysis

3.1 Vocabulary Coverage

We evaluated the synonym networks generated from both WordNet and the LLM (Mistral-small) using standard graph metrics. The LLM network contains 2,668 nodes and 5,139 edges, while WordNet includes 2,122 nodes and 2,477 edges. The higher density of the LLM network (0.00144 vs. 0.00110) indicates greater connectivity and a richer lexical reach. The LLM and WordNet networks exhibit notable differences in basic properties:

3.2 Network Topology Comparison

To evaluate the networks comprehensively, we examined three key aspects: overall scale and connectivity, structural compactness, and cohesion. After previously comparing the number of nodes and edges, we further examined both networks through structural metrics such as clustering coefficient, path length, diameter, connected components, and betweenness centrality (see Table 1).

To go beyond just counting words and connections, we examined the structure of both networks using a range of graph metrics that reveal how synonyms are organized (see Table 1). The LLM network clearly covers more ground, with more nodes (2,668 vs. 2,122) and nearly twice as many edges (5,139 vs. 2,477) as WordNet. This also makes it more densely connected (0.00144 vs. 0.00110), and with a higher average degree (3.85 vs. 2.33), each word in the LLM network tends to be linked to more synonyms.

Despite this, both networks are similarly cohesive on a local level, as shown by their nearly equal clustering coefficients. But what really sets them apart is how connected and navigable the overall structure is. Words in the LLM graph are much closer to each other—on average, it takes only 7 steps to get from one word to another, compared to 10 in WordNet. The LLM also has a smaller diameter (15 vs. 24), meaning its furthest-apart words are still relatively close.

Perhaps the most striking difference is how fragmented WordNet is, breaking into 211 disconnected parts. The LLM network, in contrast, is almost completely connected, with only 2 components. It also has higher betweenness centrality, meaning more words act as bridges that connect different areas of meaning. Taken together, these findings suggest that while WordNet is carefully structured around clear-cut word senses, the LLM network forms a much more fluid and interconnected web of language—better reflecting how words relate in natural use.

Table 1: Comparison of Graph Metrics for WordNet and LLM Synonym Networks

Metric	**WordNet**	**LLM**
Number of Nodes	2,122	2,668
Number of Edges	2,477	5,139
Density	0.00110	0.00144
Average Degree	2.33	3.85
Avg. Clustering Coefficient	0.1935	0.1917
Avg. Shortest Path Length	10.01	7.04
Diameter	24	15
Connected Components	211	2
Avg. Betweenness Centrality	0.0011	0.0022

3.3 Degree Distribution Analysis

To understand how synonym connections are distributed across words, we analyzed the degree distribution of both networks. Figure 1 shows a log-log plot of node degrees against their frequency. This illustrates how many words in each network have a certain number of synonyms.

The LLM network exhibits a broader and more gradual decay, indicating the presence of many high-degree nodes—words connected to a large number of other words. In contrast, the WordNet curve drops off sharply, showing that most words are connected to only a few others. This reflects WordNet's sense-based, manually filtered design, which prioritizes semantic accuracy and avoids overgeneralization.

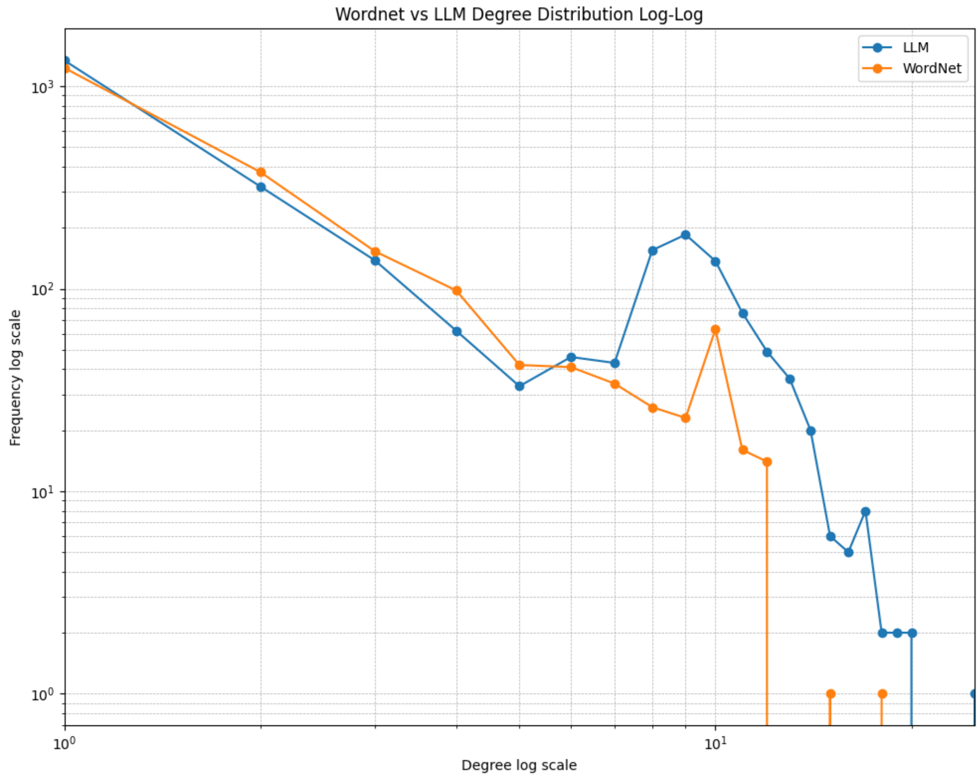

Figure 1: Degree distribution of WordNet and LLM networks on a log-log scale.

3.4 Case Study: ”heavy”

To illustrate the qualitative differences between the networks, we visualized the synonym neighborhoods of the word *heavy* within a two-hop radius. The contrast is striking. The LLM-generated network (on the left) expands across multiple semantic fields, capturing literal (*hefty*, *overweight*), abstract (*significant*, *important*), emotional (*gloomy*, *dismal*), and even evaluative terms (*stupid*, *boring*). This highlights the LLM's tendency to encode broad, polysemous relationships, consistent with prior findings on contextual word generalization (Blevins & Zettlemoyer, 2020)

In contrast, the WordNet network (on the right) remains narrower and more defined. Its connections are limited to core senses of *heavy*, such as physical weight, difficulty, and density. This focused structure reflects WordNet's manually curated, sense-specific design (Fellbaum, 1998), prioritizing semantic clarity over breadth.

Together, these graphs reinforce the idea that LLMs model how language is used in diverse, overlapping contexts, while WordNet adheres to a more categorical view of meaning.

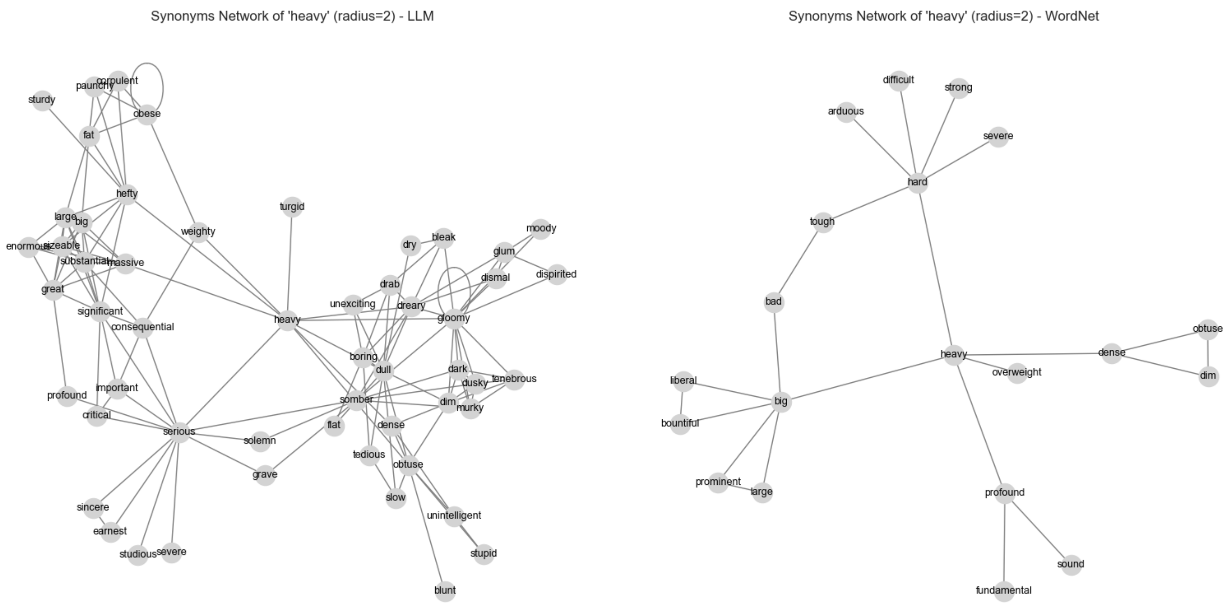

Figure 2: Comparison of synonym neighborhoods of *heavy* in LLM and WordNet networks.

4 Discussion and Conclusions

Our analysis reveals that LLM-generated synonym networks are broader and more interconnected, capturing nuanced, context-dependent associations that go beyond literal meanings (Petroni et al., 2019). In contrast, WordNet maintains a conservative and segmented structure, reflecting its expert-curated design focused on semantic precision (Fellbaum, 1998). These differences suggest that LLMs are better suited for generative, adaptive NLP tasks, while WordNet remains ideal for interpretable and controlled applications. Future research could explore hybrid systems that combine WordNet's lexical rigor with the contextual flexibility of LLMs (Camacho-Collados & Pilehvar, 2018).

References

Blevins, T., & Zettlemoyer, L. (2020). Moving beyond the word: Subword-level compositional meaning. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 5709–5721).

Camacho-Collados, J., & Pilehvar, M. T. (2018). From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, *63*, 743–788.

Fellbaum, C. (Ed.). (1998). *Wordnet: An electronic lexical database*. Cambridge, MA: MIT Press.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space.* arXiv preprint arXiv:1301.3781.

Petroni, F., Rocktäschel, T., Lewis, P., Bakhtin, A., Wu, Y., & Riedel, S. (2019). Language models as knowledge bases? In *Proceedings of the 2019 conference on empirical methods in natural language processing (emnlp)* (pp. 2463–2473).

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41–78.