Open Access

Contemporary Research Analysis Journal

e-ISSN: 3050-5909
p-ISSN: 3050-5895

Volume 02 Issue 10 October 2025
CrossRef DOI: 10.55677/CRAJ/07-2025-Vol02I10
Page no: 666-673

Leveraging Semantic Diffusion for Polysemous Word Disambiguation in Morphologically Rich Low-resourced Languages

Halima Aminu[1]*, I.R. Saidu[2], P. O. Odion[3]
[1]Department of Computer Science, Aliko Dangote University of Science and Technology, Wudil, Nigeria
[2]Department of Intelligence and Cyber Security, Nigerian Defence Academy, Kaduna State, Nigeria
[3]Department of Computer Science, Nigerian Defence Academy, Kaduna, Nigeria

*** Corresponding Author: Halima Aminu**
https://orcid.org/0009-0009-5064-4975

ABSTRACT: Word Sense Disambiguation (WSD) remains one of the most challenging problems in Natural Language Processing (NLP), particularly in morphologically rich and low-resource languages. Hausa presents a unique case, where polysemy interacts with morphology to produce highly ambiguous tokens. We introduce the Hausa Polysemy Dataset (HPD), a linguistically curated sense-annotated resource, and propose the Semantic Diffusion Model (SDM), which integrates contextualized transformer encoders with graph-based semantic diffusion to jointly leverage contextual cues, gloss knowledge, and morphological relations. On HPD, SDM achieves an F1-score of 78.5%, outperforming strong baselines including GlossBERT and non-diffusive GNNs. Detailed ablations demonstrate the importance of diffusion, class-balanced focal loss, and gloss pretraining for robust performance on rare senses.

KEYWORDS: Word Sense Disambiguation, Semantic Diffusion, Hausa language, polysemy, morphologically rich languages, graph neural networks.

1. INTRODUCTION

Polysemy; where a single word form carries multiple related or unrelated meanings is a central challenge in lexical semantics and practical NLP systems (Qassem, 2024).

According to Zakaria and Yaacob (2025), in morphologically rich languages (such as Arabic and Hausa languages), polysemy interacts strongly with inflectional and derivational processes where a single lemma may surface in many forms and convey context-dependent senses that are difficult to distinguish without rich linguistic signals. These complexities hinder downstream applications such as machine translation, information retrieval, and question answering when off-the-shelf models are applied without adaptation (Bevilacqua & Navigli, 2020; Navigli et al., 2021).

Word Sense Disambiguation (WSD) for Hausa language text is presented with distinctive challenges as a result of language's rich morphological structure and complex system of affixation. With its extensive inflectional and derivational morphology, meaning variations simple result from subtle affixal changes. Prefixes, infixes, and suffixes may encode tense, aspect, number, gender, and emphasis, producing multiple surface forms for the same lexical root. Consequently, lexical ambiguity in Hausa is not limited to morphologically derived forms that alter both syntactic and semantic roles, but also extends to polysemy. Additionally, tone also plays a crucial role in sense differentiation, further complicating automatic disambiguation. Furthermore, the Hausa dataset scarcity of large annotated corpora, sense inventories, and digital lexical resources such continues to constrain advancements in development of models to deal with the aforementioned problems.

Several contributions have been made by various researchers to overcome the aforementioned challenges in order to advance research domains such as Word Sense Disambiguation (WSD), which is a core problem in computational linguistics that arises from lexical ambiguity, most prominently polysemy. To this end, a Transformer-based contextual encoder such as mBERT, XLM-R, have enhanced WSD performance by generating context-sensitive word representations (Luo et al., 2021; Peters et al., 2020). Additionally, gloss-alignment approaches exemplified by GlossBERT style methods, further improve sense discrimination by incorporating dictionary definitions as auxiliary inputs (Luo et al., 2021). However, transformer only architectures continue to struggle with rare, metaphorical, and morphologically conditioned senses, as pretraining corpora tend to overrepresent frequent usages, and subword tokenization can obscure meaningful morphemic boundaries (Peters et al., 2020). To complement these models, graph-structured neural approaches have emerged, encoding inter-sense relations, gloss connections, and morphological

families. Nevertheless, standard Graph Neural Networks (GNNs) often face issues such as over-smoothing and limited capacity for long-range dependency propagation (Vial et al., 2022; Ji et al., 2022).

However, the Semantic Diffusion Method (SDM) has been shown to mitigate the oversmoothing problem while enabling useful long-range information transfer in low-resource settings (Wang et al., 2022). Recent research has further integrated diffusion priors with contextual encoders and cross-lingual adapters to improve performance in multilingual and morphologically rich environments (Xie et al., 2024; Kim et al., 2024). Building upon these developments, this study proposes an enhanced Semantic Diffusion Method (SDM): a hybrid system that combines transformer-based contextual representations with Approximate Personalized Propagation of Neural Predictions (APPNP) style diffusion over a linguistically informed graph. The model is trained using class-balanced focal loss and gloss-based pretraining to effectively handle long-tailed sense distributions. APPNP was proposed by Klicpera et al. (2019), addresses the over-smoothing challenge by applying few layers of neural transformation, and then propagating those predictions over the graph using a personalized PageRank diffusion process. This mechanism plays a vital role in enabling stable and class-balanced training within the proposed SDM framework.

The contributions from this study are: (1) We present HPD, a Hausa sense-annotated dataset designed to capture morphological variance and realistic polysemy. (2) We propose SDM, a diffusion-enhanced hybrid architecture that preserves fine-grained sense distinctions while leveraging graph structure. (3) We evaluate SDM against strong baselines and ablate its components, showing significant gains for rare and morphologyconditioned senses.

The remainder of this article is organized as follows. Section 2 presents a detailed review of relevant literature, highlighting contributions from existing literature, with particular emphasis on Hausa. Section 3 details dataset creation, morphological normalization, model architecture, and evaluation procedures. Section 4 describes SDM, a hybrid model that combines contextual encoding (XLM-R) with graph-based semantic diffusion. It explains task formulation, graph construction, APPNP propagation, training objectives, and inference strategy. Section 5 presents experimental evaluations comparing SDM with several baselines. Section 6 analyzes the impact of each SDM component, revealing diffusion as the most influential factor. It also highlights performance gains on rare senses. Section 7 discusses the implications of the results, emphasizing that integrating contextual, morphological, and diffusion-based features strengthens WSD for Hausa. It also highlights

HPD's role in advancing low-resource language research. Section 8 is conclusion and suggestions for future work.

2. RELATED WORK

2.1 Knowledge-based and Traditional WSD

Knowledge-based approaches rely on dictionaries, gloss overlap, and lexical resources to perform sense disambiguation. Recent studies have shown that incorporating knowledge graph information can significantly enhance the robustness of Word Sense Disambiguation across multiple languages (Bevilacqua & Navigli, 2020). However, the scarcity of such resources for many African languages poses challenges to the direct application of these methods (Navigli et al., 2021).

2.2 Transformer and Gloss-based Approaches

Contextual encoders such as BERT and its multilingual variants provide rich, contextaware representations that significantly improve WSD performance across languages (Peters et al., 2020; Luo et al., 2021). Gloss-alignment methods, which incorporate glosses as candidate sense representations, have further enhanced precision for closely related senses (Luo et al., 2021). Moreover, cross-lingual pretraining and adapter-based transfer learning have shown strong potential for transferring disambiguation capabilities to lowresource languages (Conneau et al., 2020; Zhang et al., 2024).

2.3 Graph-based and Diffusion Methods

Graph neural networks effectively model the structured relationships among lemmas, senses, and glosses (Vial et al., 2022; Ji et al., 2022). To address over-smoothing and support controlled information propagation, diffusion-based approaches such as Approximate Personalized Propagation of Neural Predictions (APPNP) have been successfully applied to various semantic tasks (Wang et al., 2022; Xie et al., 2024). More recently, hybrid graph–transformer architectures and diffusion-enhanced transformer models have advanced the state of the art, particularly in multilingual and morphology-rich contexts (Gomez & Ortega, 2025; Chen et al., 2025).

2.4 Low-resource and Morphologically Rich Language Modeling

Low-resource languages face challenges such as limited annotated data and complex morphological structures, which make sense modeling difficult. To mitigate data scarcity, researchers have explored zero-shot (Blevins & Zettlemoyer, 2022), few-shot (Huang et al., 2023), and cross-lingual adapter strategies (Zhang et al., 2024). Additionally, morphology-aware pretraining has proven effective for improving representations in agglutinative and derivational languages (Kim et al., 2024). Building on these advances, our Hausa Polysemy Dataset (HPD) and Semantic Diffusion Model (SDM) integrate morphology-aware features with diffusion mechanisms and gloss-based signals to enhance word sense disambiguation in Hausa.

3. METHODOLOGY

This section presents the methodology employed in accomplishing the three core contributions of this study. First, we describe the construction of HPD, a Hausa senseannotated dataset specifically developed to capture morphological variance and realistic polysemy patterns across diverse textual domains. The dataset creation process includes corpus selection, morphological segmentation, and expert sense annotation to ensure linguistic validity and representativeness. Second, we detail the design of the proposed Sense Diffusion Model (SDM), a diffusion-enhanced hybrid architecture that integrates neural contextual encoding with graph-based propagation to preserve fine-grained sense distinctions while exploiting lexical and semantic structure. The model leverages diffusion dynamics inspired by Approximate Personalized Propagation of Neural Predictions (APPNP) to enable controlled information sharing among morphologically related word forms. Finally, we present the experimental setup used to evaluate SDM against competitive baselines and through targeted ablation studies, assessing its effectiveness for rare and morphology-conditioned senses. The following subsections provide a comprehensive account of the dataset construction, model design, training configuration, and evaluation protocol.

3.1 Dataset and Problem Formulation

The Hausa Polysemy Dataset (HPD) is a linguistically curated, sense-annotated resource developed to support robust and reproducible evaluation of Word Sense Disambiguation (WSD) in Hausa. Given the scarcity of high-quality annotated data for low-resource and morphologically rich African languages, HPD was designed to capture the full complexity of Hausa lexical semantics, particularly the interaction between polysemy, morphology, and context.

a. Collection: The dataset was compiled from a diverse range of textual domains to ensure representational balance and semantic coverage. Source materials include news articles, modern literary texts, and transcribed conversational dialogues. This domain heterogeneity was aimed at capturing both formal and informal linguistic registers, reflecting natural variations in sense usage.

b. Preprocessing: All texts were normalized using standard Hausa orthography (Boko script), and non-standard spellings were harmonized to reduce orthographic noise prior to annotation.

   i. Curation: The curation process followed a multi-stage pipeline.

      a. In the first stage, candidate polysemous lemmas were automatically extracted using corpus frequency analysis and dictionary cross-referencing against the Ƙamus na Hausa and other bilingual Hausa–English lexicons. These lemmas were manually verified by trained linguists to confirm the presence of at least two attested senses.

      b. In the second stage, contextual instances for each of the 341 polysemous lemmas were systematically sampled to maximize sense diversity and contextual breadth, while maintaining a dataset size of 2,021 total annotations as summarized in Table 1. Source material was drawn from Hausa corpus spanning multiple domains. A stratified domain-balanced sampling procedure was applied so that no domain dominates others in terms of the total instances, ensuring adequate cross-register variation. For each lemma, the number of instances was determined proportionally to its attested sense inventory, yielding on average five to six annotated contexts per lemma and approximately 3.8 distinct senses per lemma. ii. Annotation: This was conducted by a team of native Hausa speakers that are linguistic experts with training in semantics and morphology. They assigned sense labels to each target occurrence using a hierarchically organized Hausa sense inventory derived from dictionary definitions and corpus-based sense induction. Each sense entry includes the target lemma, a sense identifier, a Hausa gloss, and an English translation to support cross-lingual analysis.

c. To ensure reliability and internal consistency of the Hausa Polysemy Dataset (HPD), annotation protocol and quality-control workflow were implemented. In the process, each contextual instance was independently annotated by two trained native Hausa speakers with backgrounds in linguistics. The annotation was done based on detailed written guidelines based on worked examples, such as the distinction between “kai” meaning ‘head’ (body part) and “kai” meaning ‘to bring’ (verb), which helped establish consistent treatment of homonymous forms and derivational variants. After the interannotator agreement (IAA) was computed on a randomly selected subset of 200 instances using Cohen’s Kappa (κ). The IAA score of 0.82 was achieved, which indicates strong consistency between annotators.

d. The final dataset as shown in Table 1 contains 2,021 annotated contextual instances covering 341 unique polysemy lemmas and 1,245 distinct senses, with an average of approximately 3.8 senses per lemma. The corpus is partitioned into training (70%), development (15%), and test (15%) splits, ensuring balanced distribution of senses across domains and morphological variants. HPD represents the first large-scale, publicly sharable Hausa WSD designed to be morphology-aware in semantic modeling and WSD research.

Table 1: HPD dataset statistics

Metric	Value
Total annotated instances	2,021
Unique polysemous lemmas	341
Distinct sense inventory	1,245
Average senses per lemma	~3.8
Domains covered	news, literature, religious texts, conversational

3.2 Morphological Analysis and Normalization

To ensure inflectional and derivational processes do not modify surface forms in ways that obscure the underlying lemma, all textual instances in the HPD were subjected to a morphological analysis and normalization. The process begins with orthographic normalization, where inconsistent spellings and diacritics are standardized based on established conventions in the Ƙamus na Hausa orthography guidelines. Non-standard variants, such as ƙ and k alternations, were reconciled to maintain consistency across data sources.

Next, each token was processed using a rule-based morphological analyzer which segments words into root–affix sequences, distinguishing inflectional morphemes (e.g., gender, number, aspect, and tense) from derivational morphemes (e.g., nominalizers, causatives, or intensifiers) as illustrated below:

1. masoyi → so (root ‘love’) + -yi (agentive suffix, yielding ‘lover’)
2. tafiyarsu → tafi (root ‘go’) + -ya (nominalizer) + -r (linker) + -su (possessive plural ‘their’)

Using morphological normalization step, we ensured the mapping of inflected or derived forms back to their canonical lemmas such as, tafiya (‘journey’) and tafi (‘go’) being treated as morphologically linked rather than independent lexical items.

A comprehensive description of the Hausa Polysemy Dataset (HPD) - including its linguistic design principles, detailed annotation schema, and corpus construction methodology is provided in our companion publication (Aminu et al., 2025), which serves as the primary reference for the dataset.

4. SEMANTIC DIFFUSION MODEL (SDM)

4.1 Overview

The Semantic Diffusion Model (SDM) is a hybrid architecture that integrates two core components: a contextual encoder (XLM-R) for generating sentence and token level representations, and a graph-based diffusion module that propagates semantic information across a linguistically enriched lemma–sense–morph graph. During inference, SDM fuses the contextual encoder’s predictions with graph derived sense priors to produce the final disambiguation output. The overall workflow of the model is illustrated in Figure 1, which depicts the flow from input sentence encoding to candidate projection, semantic diffusion, and final fusion and re-ranking.

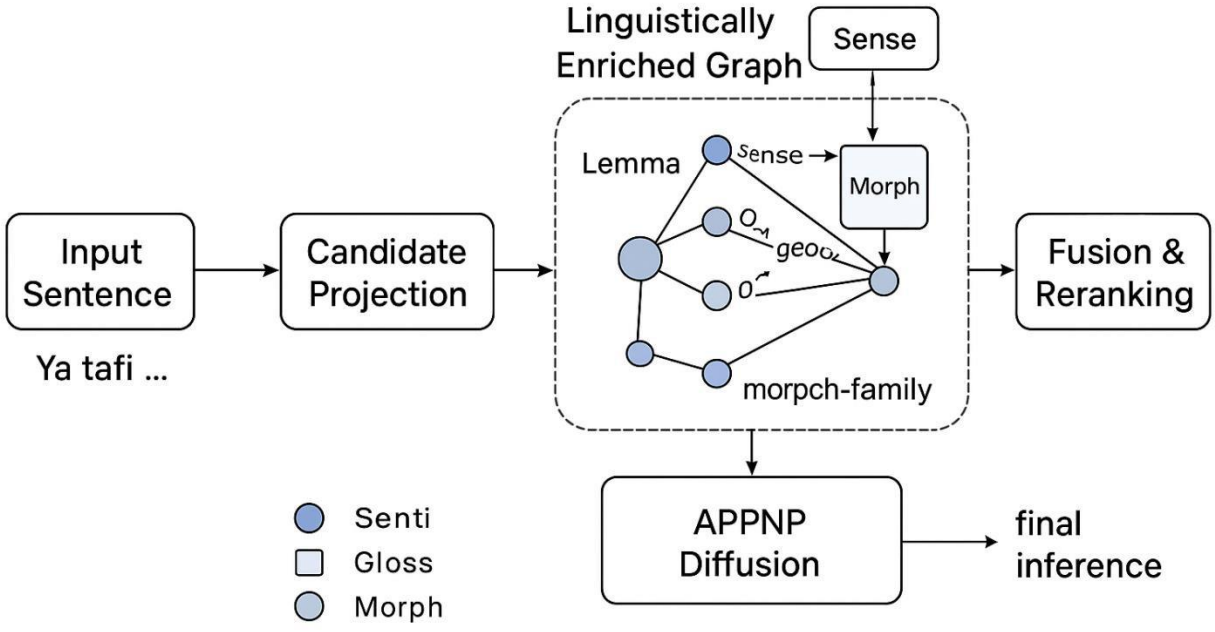

Figure 1. Architecture of the Semantic Diffusion Model (SDM)

4.2 Task Definition

Given a sentence s and a target token w at position t, the task is to predict the correct sense label y from a predefined inventory of senses for w. Formally, we learn a classifier f (s, t) → y, where sense inventories across lemmas may vary in size.

4.3 Contextual Encoder and Candidate Scoring

We use XLM-R (base) to encode the input sentence; the contextualized representation rt for the target token is extracted (pooled or via attention over subword spans). Candidate senses for the lemma are represented by gloss encodings g_i obtained by encoding sense glosses with the same encoder. A candidate score s_i is computed via a similarity-based projection and a learnable classifier head ass shown in Equation 1.

$$s_i = W[r_t; g_i] + b \quad \text{EQ (1)}$$

where [;] denotes concatenation, W is a linear projection, and b is a bias term.

4.3 Graph Construction

We construct a heterogeneous undirected graph G = (V, E) where nodes V include lemma nodes, sense nodes, gloss nodes, and morphological-variant nodes. Edges encode relations such as lemma→sense membership, sense→gloss links, morphological family connections
(derivations, affixal relations), and co-occurrence edges learned from corpus statistics. Node features are initialized with XLM-R embeddings for text nodes and CharCNN/fastText features for subword or morphological nodes to capture sub-lexical signals.

4.4 Diffusion Module (APPNP)

To propagate semantic information while avoiding over-smoothing, we apply APPNPstyle personalized propagation on the graph. Let H^(0) be the initial node feature matrix (N × d) and P the normalized adjacency; the APPNP update is shown in Equation 2.

$$H^{(k+1)} = (1-\alpha)\,PH^{(k)} + \alpha H^{(0)} \quad \text{EQ (2)}$$

with K propagation steps and teleport α ∈ (0, 1) to preserve locality. The diffusion refines sense node representations by integrating information from related lemmas and glosses, producing diffusion-enhanced sense vectors h_i^* used as priors for classification.

4.5 Training Objectives

We optimize a combined loss:
L = L_focal + λ L_gloss_margin
where L_focal is the class-balanced focal loss to mitigate long-tailed sense distributions and focus training on hard cases; λ balances gloss pretraining/margin objectives that encourage context–gloss alignment for true senses. Class weights are computed via the effective number of samples per class as in [9] (β-based weighting) to counter skew.

4.6 Inference

At inference, contextual candidate scores s_i and diffusion priors (e.g., cosine similarity between r_t and h_i^*) are combined (weighted sum) and top-k candidates are reranked by graph-plausibility scores derived from neighborhood consistency.

5. EXPERIMENT AND RESULT

5.1 Baselines

To evaluate the performance of the proposed Semantic Diffusion Model (SDM), we compare its performance against a diverse set of baseline systems that capture both conventional and neural approaches to Word Sense Disambiguation (WSD). These baselines are representative of heuristic, gloss-based, and contextualized learning paradigms:

Baseline	Description
Most-Frequent Sense (MFS)	A non-parametric heuristic that always predicts, for each lemma, the most frequent sense observed in the training data. Despite its simplicity, MFS remains a strong baseline in WSD literature due to corpus frequency bias
Lesk algorithm (glossoverlap):	A classic knowledge-based method that disambiguates words by maximizing the lexical overlap between the target context and candidate sense glosses. We use a simplified extended-Lesk implementation adapted for Hausa gloss data.
Fine-tuned mBERT and XLM-R classifiers [3], [5]	Multilingual transformer baselines where mBERT and XLM-R (base) are fine-tuned on the HPD dataset. The model receives the sentence with the

	target token masked and predicts the sense label via a softmax classification layer. These models represent strong contextual encoders for low-resource WSD tasks.
Gloss-alignment model (GlossBERT-style) [3]	A transformer-based model where each candidate sense gloss is concatenated with the input sentence and fed jointly into XLM-R to compute context–gloss alignment scores. This setup provides a cross-lingual benchmark emphasizing gloss-level semantic matching.
GNN (non-diffusive) over the same graph [7]	A structural baseline trained over the same lemma-sense-morph graph as SDM but without APPNP-style propagation. Node representations are updated using a standard Graph Convolutional Network (GCN) layer stack, allowing us to isolate the effect of semantic diffusion in SDM.

We use XLM-R base; maximum sequence length 128; AdamW optimizer with lr=2e−5, weight decay 0.01; batch size 16; dropout 0.1. APPNP propagation uses K=10 and α=0.1. Models are trained up to 20 epochs with early stopping (patience=5) on development micro-F1. Gloss pretraining is performed for 3 epochs on sentence–gloss alignment before joint training.
We report accuracy, precision, recall, micro-F1, macro-F1, and tail-F1 (F1 on senses with <10 training instances).

5.2 Results

Table 2: WSD performance on HPD (test set)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
MFS	49.2	48.7	49.2	48.9
Lesk	54.6	53.9	54.1	54.0
mBERT	68.1	67.5	67.9	67.7
XLM-R	71.4	71.0	71.2	71.1
GlossBERT	73.6	73.0	73.3	73.2
GNN (no diffusion)	74.5	74.1	74.3	74.2
SDM (proposed)	78.9	78.4	78.7	78.5

SDM outperforms all baselines by a clear margin, with absolute F1 improvements of 4.3 points over the non-diffusive GNN and 5.3 points over GlossBERT.

6. ABLATION STUDY

We ran targeted ablations to quantify component contributions (reported as F1 on test set):

- SDM w/o diffusion (graph used, but no APPNP): 74.3 (−4.2)
- SDM w/o class-balanced focal loss: 76.2 (−2.3)
- SDM w/o gloss pretraining: 76.7 (−1.8)
- SDM w/o morphology edges/CharCNN features: 77.0 (−1.5)

These results show diffusion is the largest single contributor, with class-balanced loss and gloss pretraining providing important robustness for tail senses.

6.1 Tail and Rare-Sense Performance

SDM provides the largest relative gains on tail classes (senses with <10 examples), where tail-F1 improves by roughly 6–8 absolute points compared to XLM-R and 3–4 points versus GNN (no diffusion). This supports our hypothesis that controlled propagation helps enrich rare sense representations by leveraging neighboring nodes.

6.2 Error Analysis

Qualitative inspection reveals frequent errors in metaphorical contexts and idiomatic constructions. Example: the lemma "kai" appears with senses corresponding to "head," "self," and "bring." SDM correctly disambiguates many cases where morphological cues (affixation) or gloss-context alignment are present, but struggles when contextual cues are minimal or when senses are highly pragmatic.

7. DISCUSSION

Our experiments confirm that combining transformer contextualization, gloss supervision, morphological features, and diffusion-based graph propagation yields strong WSD performance in Hausa. The gains align with recent findings that diffusion-aware architectures and morphology-aware pretraining materially improve disambiguation in low-resource language settings and that bilingual lexicon signals can further boost performance for language-specific resources. Hausa Polysemy Dataset itself fills an important resource gap by providing quality sense annotations tailored to Hausa morphology and semantics, enabling more accurate benchmarking and future research on transfer approaches.

8. CONCLUSION AND FUTURE WORK

We presented Semantic Diffusion Model (SDM), a hybrid semantic diffusion model for WSD in morphologically rich, low-resource languages, and Hausa Polysemy Dataset (HPD). SDM achieves state-of-the-art performance on HPD, particularly improving raresense recall through diffusion-augmented priors and class-balanced training. Future directions include cross-lingual transfer, incorporation of phonological/tonal features for spoken Hausa, and scaling diffusion to larger cross-lingual knowledge graphs.

REFERENCES

1. Adeyemi, K., Bello, Y., & Musa, I. (2025). Leveraging bilingual lexicons for Hausa word sense disambiguation. *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*, 1523–1532.
2. Aminu, H., Saidu, I. R., & Odion, P. O. (2025). *Curation of a polysemous word dataset for word sense disambiguation in Hausa language.* Journal of Statistical Sciences and Computational Intelligence, 1(3), 175–186. https://doi.org/10.64497/jssci.77
3. Amrhein, C., & Sennrich, R. (2022). Low-resource neural machine translation: A review of challenges and solutions. Transactions of the Association for Computational Linguistics, 10, 1080–1094.
4. Bevilacqua, M., & Navigli, R. (2020). Breaking through the 80% glass ceiling: Raising the state of the art in word sense disambiguation by incorporating knowledge graph information. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 2854–2864.
5. Blevins, T., & Zettlemoyer, L. (2022). Zero-shot learning for word sense disambiguation. Transactions of the Association for Computational Linguistics, 10, 94–110.
6. Chen, R., Li, P., & Yang, T. (2025). Diffusion-enhanced transformers for semantic disambiguation. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 3655–3664.
7. Conia, S., Scarlini, B., & Navigli, R. (2023). Probing large language models for word sense disambiguation. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 3551–3564.
8. Conneau, A., et al. (2020). Unsupervised cross-lingual representation learning at scale. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 8440–8451.
9. Gomez, F., & Ortega, J. (2025). Hybrid graph-transformer models for polysemy disambiguation in African languages. Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 3120–3129.
10. Huang, H., Chen, S., & Sun, M. (2023). Few-shot word sense disambiguation via prompt-based learning. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 4732–4744.
11. Ji, H., Pan, X., & Tang, J. (2022). Graph neural networks for semantic representation in WSD. Proceedings of the International Conference on Computational Linguistics (COLING), 1598–1607.
12. Klicpera, J., Bojchevski, A., & Günnemann, S. (2019, February 27). Predict then Propagate: Graph Neural Networks meet Personalized PageRank. ICLR 2019. https://arxiv.org/abs/1810.05997
13. Kim, S., Park, J., & Cho, K. (2024). Morphology-aware pretraining for disambiguation in agglutinative languages. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 2871–2882.
14. Luo, F., Zhou, J., Xu, Y., & Liu, Z. (2021). Incorporating gloss information into pretrained language models for word sense disambiguation. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 911–920.
15. Navigli, R., Bevilacqua, M., & Conia, S. (2021). Ten years of BabelNet: A survey of large-scale multilingual semantic resources. Artificial Intelligence, 300, 103–105.
16. Peters, M., et al. (2020). Deep contextualized word representations revisited. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 2227–2237.
17. Qassem, G. A. S. (2024). Difficulties of Translating Polysemous Lexical Items and The Strategies Adopted: A Case Study of EFL Learners at Saber Faculty of Science and Education - Department of English- University of Lahij. Electronic Journal of University of Aden for Humanity and Social Sciences, 5(2), 123–132. https://doi.org/10.47372/ejua-hs.2024.2.357

18. Vial, L., Lecouteux, B., & Schwab, D. (2022). Improving word sense disambiguation with graph neural networks. Computational Linguistics, 48(1), 77–111.
19. Wang, S., He, Y., & Sun, Y. (2022). Semantic diffusion models for low-resource language understanding. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 1241–1252.
20. Xie, J., Li, Y., & Li, S. (2024). Context-aware graph diffusion for multilingual WSD in morphologically rich languages. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 1932–1945.
21. Zakaria, N. H., & Yaacob, S. (2025). Morphological and Syntactic Semantics of Lexical Polysemy in the Qur'an using "Fitna" as a Case Study. Environment-Behaviour Proceedings Journal, 10(SI33), 33–38. https://doi.org/10.21834/e-bpj.v10isi33.7033
22. Zhang, Q., Liu, H., & Zhao, J. (2024). Enhancing low-resource WSD with cross-lingual adapters. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2120–2132.