

Minimax Optimization for Ambiguity Resolution in Multilingual Text Analysis

Kamolbek Masharipov¹, Zarif Zairov², Mashkhura Husanova³, Dilfuza
Khudayberdiyeva⁴, Ilkhom Kholmuminov⁵, Guzal Davurova⁶, Umida Saribaeva⁷

¹Deputy Dean, Mamun University, Khiva, Uzbekistan. Email: lkamolbek253@mamunedu.uz,
<https://orcid.org/0009-0003-2002-7534>

²Associate Professor, Department of Music Education, Termez State University, Termez,
Uzbekistan.

Email: zairov@tersu.uz, <https://orcid.org/0000-0001-9145-1787>

³Teacher, Department of English Philology and Teaching Methodology, Andijan State
Institute of Foreign Languages, Andijan, Uzbekistan. Email: mashxura969@gmail.com,
<https://orcid.org/0009-0002-0807-0911>

⁴Associate Professor, Samarkand State Institute of Foreign Languages, Samarkand,
Uzbekistan.

Email: hudayberdiyevadilfuza3@gmail.com, <https://orcid.org/0009-0000-4145-0938>

⁵Associate Professor, HOD, Department of Foreign Philology, Tashkent International
University of Financial Management and Technologies, Tashkent, Uzbekistan. Email:
ilxomxolmuminov713@gmail.com, <https://orcid.org/0000-0002-2904-7879>

⁶Teacher, Department of Pedagogy, Psychology and Inclusive Education, Jizzakh State
Pedagogical University, Jizzakh, Uzbekistan. Email: guzaldovurova@gmail.com,
<https://orcid.org/0000-0002-0082-8583>

⁷Associate Professor, Department of Psychology, Candidate of Psychological Sciences,
University of Tashkent for Applied Science, Tashkent, Uzbekistan. Email:
umidasaribayevna@gmail.com, <https://orcid.org/0009-0005-7230-7313>

Abstract

Multilingual text analysis needs a procedure capable of differentiating between interpretations based on their dependency on context, language, and presence of lexical information. Selecting the interpretation with maximal probability can result in a decision error under circumstances of uncertain probabilities. In this paper we propose a minimax decision procedure for solving the problem of lexical ambiguity resolution. The contextual module generates interpretations and associated nominal probabilities; an uncertainty set is defined for a set of alternative probability distributions. In the decision module the interpretation with the minimum worst-case expected loss is chosen or a decision of deferment is made if the cost of a decision error outweighs the cost of deferral. For binary ambiguities, our approach boils down to a clear-cut criterion using the extremes of the interval probability distribution. We present a numerical example with six cases in English, Spanish, and Romanized Hindi, comparing the forced prediction, threshold-based selection, and minimax selection approaches. Based on our assumptions regarding probabilities and losses, minimax selection resolves three out of six cases and cuts the average worst-case expected decision loss from 0.4383 to 0.2250. Its nominal expected loss rises to 0.1867, starting from 0.1717 due to the deferral penalty for the unassigned items. The above examples show the tradeoff between robustness and coverage that cannot be described by a language benchmark. The method disentangles the uncertainty over probabilities from the uncertainty in the text and allows for inspectable decision-making. The future assessment needs to derive the uncertainty sets based on multilingual validation corpora.

Keywords: Multilingual NLP, lexical ambiguity, minimax decision making, word sense disambiguation, selective prediction.

1. Introduction

There can be various interpretations of a particular word, and the context may not provide sufficient evidence for picking an interpretation. The process of disambiguation helps solve this problem since it resolves the meaning of the word based on the context (Navigli, 2009). In the case of multilingual disambiguation, the difficulty of

ambiguity is increased because of the different grammatical, lexical, writing, and conceptual structures of the languages. An ambiguity in English doesn't necessarily exist in other languages.

The analysis system must then distinguish between cases in which the word “bank” means either “financial institution” or “river’s edge,” “banco” means either “financial institution” or “bench” in Spanish, or “kal” means either “yesterday” or “tomorrow” in Hindi. Context can clear up the ambiguity, but when dealing with fragmented text, missing time-related information, and translation, multiple interpretations of the terms become possible. Hence, these instances should be modeled on the basis of candidate senses in each respective language.

The use of multilingual semantic and contextual representations can be an approach to linking concepts across different languages (Navigli & Ponzetto, 2010; Pires et al., 2019; Conneau et al., 2020). Nevertheless, having a good representation does not ensure the reliability of all predicted probabilities. In cases when the data is sparse or the language is under-represented, a high probability may go hand in hand with uncertainty.

A minimax decision making level is developed which assesses a possible meaning under different probability distributions that can apply. It chooses one meaning or explicitly refuses to do so. The novelty is an easily interpretable formalization that links the problem of multilingual ambiguity with the uncertainty of probabilities and cost of misinterpretation. A specific numeric example illustrates the decision-making rules. No multilingual model has been trained and any benchmark accuracy results are not given. Multilingual XLM-R and cross-lingual learning improve low-resource language processing and help understand multilingual text representations and disambiguation problems (Gurusamy & Janarthanan, 2026). Quality evaluation of machine translation helps analyze translation accuracy, multilingual text processing, and language ambiguity in educational communication (Al-Rawajfeh, 2024).

2. Background and problem motivation

2.1. Sense inventories and contextual evidence

Word Sense Disambiguation traditionally relies on selecting possible senses and understanding context clues. Navigli's survey addresses the issues of sense differentiation and evaluation methods (Navigli, 2009). Babel Net offers a multilingual semantic network capable of helping with generation of candidates and interlingual links (Navigli & Ponzetto, 2010). This tool is quite helpful, although there may be cases when a possible sense will be missed by the candidate generator or distinguished too finely.

Raganato et al. proposed a unified framework of evaluating Word Sense Disambiguation approaches (Raganato et al., 2017). Gloss BERT provides a combination of definitions and context through the context-gloss pairs (Huang et al., 2019). These examples lead to the development of a modular framework where candidate generator and contextual scorer provide necessary information to the decision module. In our case, the decision procedure is altered, but not the sense definition.

2.2. Transfer across languages and meaning comparisons

Pires et al. studied cross-lingual transfer in multilingual BERT and observed systematic differences depending on the pairs of languages (Pires et al., 2019; Rustamov et al., 2026). Conneau et al., (2020) created XLM-R using massive multilingual pretraining. These researches have shown the importance of representation sharing, but do not solve the question of making reliable decisions about uncertain examples.

WiC is a task which aims to check whether the target word has the same meaning in two sentences (Pilehvar & Camacho-Collados, 2019). XL-WiC is a similar type of task applied to different languages Raganato et al., (2020), and MCL-WiC uses both multilingual and cross-lingual comparisons (Martelli et al., 2021). This kind of tasks provides a good evaluation direction for our future solution, but differs from choosing only one meaning from an inventory. An implementation of such a model will require changes in label candidates and loss definition. Real-time translation of multilingual medical terminology shows the necessity of disambiguation of linguistic forms and semantic correctness (Kapoor & Gupta, 2023).

2.3. Confidence uncertainty and robust decisions

Guo et al. prove that confidence estimation with neural networks requires calibration (Guo et al., 2017). Calibration deals with the connection between probabilities and frequencies of correctness; however, it does not resolve any ambiguity in a concrete sentence. The sentence can still be ambiguous although the model is calibrated on similar cases.

Robust optimization deals with decisions that are evaluated under explicitly described unfavorable conditions (Bertsimas et al., 2011). Distributionally robust learning is an extension of this view on uncertainty for data

generation distribution, while group-robust learning focuses on uncertainty in the performance of shifts affecting pre-defined groups. In this paper, the principle of decision making is applied to instances of language: uncertainties correspond to the probabilities of alternative explanations of the concrete sentence. Multilingual philology from cognitive and neuroscientific views gives some background for multilingual language processing and interpretation.

There is also another complementary approach to this problem. Specifically, Chow studied the tradeoff between recognition errors and rejection, while Geifman and El-Yaniv investigated selective classification using deep networks (Blodgett et al., 2020). The described solution is a combination of a deferral move and worst-case expected interpretation loss. This is an application of already known concepts, not invention of minimax optimization or abstention techniques.

3. Proposed method

3.1. Language-sensitive candidate construction

Let us consider x text as the textual context, t as the target phrase, and l as the language. The set $S(x, t, \ell) = \{s_1, \dots, s_K\}$ of possible interpretations is given. The multilingual encoder and the dictionary can together score how well each candidate fits into the context. Morphology and tenses of the specific language must be taken into account while scoring.

In the case of comparable ideas in other languages, the semantic knowledge base will provide the correct identifier for the meaning when possible (Navigli & Ponzetto, 2010). The alignment process should not constrain a meaning into a category that does not exist or is culturally wrong in the target language. If the meaning does not appear in the candidate generation step, there is no way for the decision layer to retrieve the meaning through probability weighting alone.

The scorer produces a nominal probability vector $\hat{p} = (\hat{p}_1, \dots, \hat{p}_K)$. These values may be calibrated using a development set Guo et al., (2017). They are inputs to the robust decision, not assumed ground truth.

3.2. Probability envelopes for uncertain interpretations

Instead of trusting a single vector, define a nonempty set of plausible distributions:

$$\mathcal{P}(x, t, \ell) = \left\{ p: p_i \geq 0, \sum_{i=1}^K p_i = 1, L_i \leq p_i \leq U_i \right\}$$

The lower and upper bounds describe uncertainty about the interpretation probabilities. Bounds must be coherent, with $\sum_i L_i \leq 1 \leq \sum_i U_i$. This condition is sufficient for a nonempty interval-constrained simplex when each lower bound does not exceed its upper bound.

One possible way to determine uncertainty would be to measure it based on out-of-sample calibration behavior, variability among model fits, or domain shift examples from literature. All constructions should be validated. Ensemble agreement does not guarantee statistical confidence and an arbitrary-wide interval can result in the system refusing good candidates.

The set denotes uncertainty about probability, while multiple interpretations denote uncertainty about meaning. These are two different things that are related. Wide intervals do not imply that all candidates are equally probable. This means that the system is not ready to consider its probabilities precise.

3.3. Interpretation losses and the deferral action

The available actions are $A = S \cup \{\perp\}$, where $\perp$ denotes abstention or referral for additional context. For the simple model, a correct interpretation costs zero, an incorrect interpretation costs one, and abstention costs c , with $0 < c < 1$:

$$\mathcal{L}(a, s_i) = \begin{cases} 0, & a = s_i, \\ 1, & a \in S, a \neq s_i \\ c, & a = \perp. \end{cases}$$

The deferral penalty represents the hassle or effort related to the nonresolution of a case. It does not mean that human review is error-free. In a practical model, review error, delay, and language-specific review effort must be accounted for. Where misunderstanding of two specific meanings causes particularly serious damage, the unit error penalty may be substituted by a specific application cost matrix.

3.4. The deterministic minimax decision

For a chosen action, worst-case expected loss is

$$R(a \mid x, t, \ell) = \max_{p \in \mathcal{P}(x, t, \ell)} \sum_{i=1}^K p_i \mathcal{L}(a, s_i)$$

The system selects

$$a^* \in \operatorname{argmin}_{a \in A} R(a \mid x, t, \ell)$$

Outer minimization uses a deterministic or evasive strategy; inner maximization uses the worst-case probability distribution consistent with the probability set. There is no assumption that speakers have antagonistic goals. Given a finite action space and a polyhedral probability space, the inner problem becomes a small linear program (Bertsimas et al., 2011). Randomized interpretation strategies are not considered in this paper.

Under zero-one interpretation loss, the risk of selecting s_j simplifies to

$$R(s_j) = 1 - \min_{p \in \mathcal{P}} p_j, \quad R(\perp) = c$$

For interval-constrained probabilities, the smallest feasible value of p_j is

$$p_j^{\min} = \max \left(L_j, 1 - \sum_{i \neq j} U_i \right)$$

The criterion is hence the selection of the alternative whose lowest attainable probability is maximum if the highest risk of the minimum probability does not surpass c ; otherwise, the action should be not to make a choice. In case of a tie between the latter and a choice, interpretation will prevail.

3.5. A closed-form rule for two competing meanings

For two candidates, write $p = P(s_1 \mid x, t, \ell)$ and let $p \in [L, U]$. The risks become

$$R(s_1) = 1 - L, \quad R(s_2) = U, \quad R(\perp) = c$$

Thus s_1 is selected when $1 - L$ is the smallest risk and is no greater than c . Candidate s_2 is selected when U is the smallest risk and is no greater than c . Otherwise, the text is deferred. This rule uses interval endpoints rather than the nominal maximum probability alone.

With $c = 0.25$, candidate s_1 requires $L \geq 0.75$, and candidate s_2 requires $U \leq 0.25$. A text with nominal probability 0.80 and a plausible interval of $[0.35, 0.90]$ can therefore be deferred despite apparently high confidence. All intervals in the worked example contain their nominal values.

4. Worked Analysis and Evaluation Plan

4.1. Three languages and six illustrative contexts

The example consists of two context types per language. For English, “The bank approved the loan” provides a financial clue, while “They discussed the bank” allows more freedom in understanding both financial and riverbank meanings. For Spanish, “Solicitó un préstamo en el banco” has a loan clue, and “Hablamos del banco” is a smaller phrase in which either a financial institution or a bench can be meant. For romanized Hindi, “Main kal aaunga” has a future verb helping the meaning of tomorrow, and “Kal ki baat” does not have a time clue at all and may refer to both yesterday and tomorrow.

The Hindi sentence is written in romanized letters for better understanding; romanization is not considered to eliminate ambiguity. Candidate s_1 is the first sense in the pairs: financial institution and riverbank for

English, financial institution and bench for Spanish, and tomorrow and yesterday for Hindi. There are no shared universal sense labels in these pairs.

Probabilities for table 1 are provided below. Narrower intervals are used for clear-cue contexts, while wider intervals are utilized in short fragments and can alter the order of interpretations. This is a manually constructed assumption and not an output of BERT, XLM-R, or any trained disambiguation model.

Table 1: Assumed probability of the first candidate interpretation.

Case	Language and context type	Nominal probability	Plausible interval
E1	English with loan cue	0.90	[0.82, 0.94]
E2	English short fragment	0.80	[0.35, 0.90]
S1	Spanish with loan cue	0.85	[0.80, 0.92]
S2	Spanish short fragment	0.78	[0.30, 0.90]
H1	Hindi with future cue	0.88	[0.78, 0.94]
H2	Hindi without tense cue	0.76	[0.32, 0.90]

4.2. Competing decision policies

The three methods are tested using identical inputs. Forced prediction chooses the most probable meaning in every situation. Confidence-threshold method chooses the meaning if its probability exceeds 0.80 and does not choose otherwise. Minimax chooses between the two worst-case scenarios considering a deferral cost of 0.25.

Minimax decision is summarized in table 2 below. All probabilities are above 0.50 and, therefore, forced prediction chooses S_1 in every case. This choice is more susceptible because of wider intervals in the short sequences.

Table 2: Worst-Case risks and minimax decisions with deferral cost 0.25.

Case	Risk of first meaning	Risk of second meaning	Minimax decision
E1	0.18	0.94	First meaning
E2	0.65	0.90	Defer
S1	0.20	0.92	First meaning
S2	0.70	0.90	Defer
H1	0.22	0.94	First meaning
H2	0.68	0.90	Defer

Nominal confidence of 0.80 can be observed in E_2 . The lowest bound of the first interpretation is 0.35; therefore, the error probability of E_2 is 0.65. Deferral costs 0.25, which makes it a better choice than E_2 . This example shows that a nominal confidence limit and an uncertainty-based criterion can diverge.

4.3. Loss reduction and the coverage tradeoff

The coverage is the proportion of cases that are interpreted. The nominal expected decision loss is determined based on the nominal probabilities. The worst-case expected decision loss is computed for the given intervals, and this incorporates the deferral cost. The cases are equally weighted, and the intervals can be different for each case.

For a policy π applied to n contexts, these summary measures are

$$\text{Coverage}(\pi) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}\{\pi(x_k) \neq \perp\}$$

$$\bar{R}(\pi) = \frac{1}{n} \sum_{k=1}^n \max_{p \in \mathcal{P}_k} \mathbb{E}_p[\mathcal{L}(\pi(x_k), s)]$$

Table 3: Calculated policy outcomes on the six constructed cases.

Policy	Resolved cases	Coverage in percent	Nominal expected loss	Worst-case expected loss
Forced prediction	6	100.00	0.1717	0.4383
Confidence threshold 0.80	4	66.67	0.1783	0.2917
Minimax with deferral	3	50.00	0.1867	0.2250

The minimax strategy achieves reduction of 48.67% of average worst-case expected decision loss compared to forced prediction, and accomplishes it not by having better classification accuracy, but by making fewer decisions. The reason for its worse nominal loss is the fact that deferment costs more than the nominal error probability on some rejected examples.

Average nominal error probability for the three cases that are accepted by the minimax rule is 0.1233 and average worst-case error probability is 0.2000. Both of these metrics are based on model expectations, and no labels have been sampled in this example, thus no confusion matrix and F1-score are possible.

Changing the penalty value means changing the amount of ambiguity that will be tolerated by the system. At penalties equal to 0.10, 0.20, 0.25, and 0.70, the rule resolves 0, 2, 3, and 6 cases, respectively. Thus, coverage varies from 0% to 100%. Different losses at different penalties will mean different things operationally and must not be compared to each other.

4.4. Benchmark validation and ablation design

The future empirical study must decide on either inventory sense selection or word-in-context pair classification. For the latter task, the unified WSD approach allows for inventory sense selection, whereas WiC, XL-WiC, and MCL-WiC allow for meaning comparison according to their task definitions (Raganato et al., 2017; Pilehvar & Camacho-Collados, 2019; Martelli et al., 2021). A pairwise approach would use “same meaning” and “different meaning” as candidate labels rather than adopting the illustrative pairs directly.

Data for training, calibration, and evaluation must be independent. Performance of candidate generation, probability calibration, coverage, selectivity, and linguistic results must be evaluated. It is crucial that a direct comparison be made in terms of coverage, otherwise one system may seem safer simply because it abstains from more inputs. Review workload and human-resolution accuracy must also be evaluated.

Ablations that could be performed include removal of the uncertainty set, removal of abstention, variation in interval width, and replacement of the multilingual encoder while holding constant the decision layer. Performance metrics for such ablations cannot be provided since the appropriate experiments have not yet been run.

5. Discussion

5.1. Robustness cannot recover missing meanings

The scheme may lead to the more conservative decision but it will not allow generating the meaning which is beyond the scope of the candidate inventory. Even a narrow distribution over the set of the incorrect candidates may lead to a confident incorrect conclusion. Candidate recall and gloss accuracy should therefore be tested before interpreting the risk bound as relevant.

The conditional guarantee is partial – in case the actual conditional interpretation distribution is from the uncertainty set, the decision loss expectation will not exceed the derived bound. The above statement is not a statistical guarantee that the distribution is from the set. Moreover, it does not guarantee that the individual prediction would be right.

5.2. Context enrichment before human referral

Abstention doesn’t have to be the end of the process. The deferral of a case may result in recalling the prior sentence, questioning the user, or consulting a person knowledgeable about the language. A different context will alter the nominal probabilities and the uncertainty set. The minimax solution will be recalculated at that point, instead of being considered final. The cost of requesting more context is that it might reveal more text. An applied algorithm must decide on a proportional course of action and whether the additional context has indeed decreased the uncertainty. The human referral should also be analyzed instead of being automatically considered a zero-error event.

5.3. Unequal language coverage and review access

A language with less data will create wider uncertainty intervals and more deferrals. While deferral can lead to decreased modeled risk exposure, it can also make the service less useful to that particular language community. Average loss minimization alone is hence not enough.

Under the current model, some important issues include who is automatically served, who has to wait for the review process, and the availability of competent reviewers. Coverage and delay by language need to be disclosed in addition to accuracy. Wider uncertainty intervals should be supported by data and not assumptions about reliability of a certain language.

5.4. Scope of the mathematical illustration

The analysis considers only the finite lexicon, determinism, zero-one interpretation loss function, and probabilistic independence. It does not resolve syntactic ambiguity, discourse reference, or code switching in general, though the decision framework can be modified to apply to these problems. The example is small and artificial, and the percentage gain is applicable to its particular numbers.

In more realistic applications, it would be necessary to assess uncertainty structure, asymmetry in the interpretation loss, noise in the review process, and changes in the domain or dialect. The set of uncertainty must be validated for under coverage and conservatism. They determine whether the robust decision framework yields a practical gain above the confidence level.

6. Conclusion

In this research, multi-lingual ambiguity resolution was set up as a minimax decision with a deferral strategy included. Potential meaning selections are compared based on reasonable probability distributions, and binary ambiguity has a clear endpoint interval rule. For example, in the case study involving six decisions, the method decreases the maximized expected decision error from 0.4383 to 0.2250 and decreases the automatic solution coverage from 100% to 50%. This reveals the cost associated with the trade-off between automatic solution and probability uncertainty prevention. Its empirical applicability should be verified using candidate selection by languages, uncertainty estimation, comparison of coverage, and review measures.

References

1. Al-Rawajfeh, A. E. (2024). Machine Translation Quality Evaluation for Multilingual Educational Communication. *International Journal of English and Education*, 13(4), 86-94. <https://doi.org/10.67050/IJEE/V13I4/IJEE244007>
2. Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM review*, 53(3), 464-501. <https://doi.org/10.1137/080734510>
3. Blodgett, S. L., Barocas, S., Daumé Iii, H., & Wallach, H. (2020, July). Language (technology) is power: A critical survey of “bias” in NLP. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 5454-5476). <https://doi.org/10.18653/v1/2020.acl-main.485>
4. Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Stoyanov, V. (2020, July). Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 8440-8451). <https://doi.org/10.18653/v1/2020.acl-main.747>
5. Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017, July). On calibration of modern neural networks. In *International conference on machine learning* (pp. 1321-1330). PMLR.
6. Gurusamy, D., & Janarthanan, M. (2026). Multilingual XLM-R for Mental Health Signal Detection in Low-Resource Tamil Language Through Cross-Lingual and Bilingual Learning. *Archives for Technical Sciences*, 2(36), 963–982. <https://doi.org/10.70102/afts.2026.1836.963>
7. Huang, L., Sun, C., Qiu, X., & Huang, X. J. (2019, November). GlossBERT: BERT for word sense disambiguation with gloss knowledge. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3509-3514). <https://doi.org/10.18653/v1/D19-1355>
8. Kapoor, A., & Gupta, R. (2023). Development of a real-time multilingual medical terminology translator for emergency settings. *Global Journal of Medical Terminology Research and Informatics*, 1(1), 16-19.

9. Martelli, F., Kalach, N., Tola, G., & Navigli, R. (2021, August). SemEval-2021 task 2: Multilingual and cross-lingual word-in-context disambiguation (MCL-WiC). In *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)* (pp. 24-36). <https://doi.org/10.18653/v1/2021.semeval-1.3>
10. Navigli, R. (2009). Word sense disambiguation: A survey. *ACM computing surveys (CSUR)*, 41(2), 1-69. <https://doi.org/10.1145/1459352.1459355>
11. Navigli, R., & Ponzetto, S. P. (2010, July). BabelNet: Building a very large multilingual semantic network. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 216-225).
12. Rajan, C. (2026). Compiler-Guided Mixed-Precision Risc-V Accelerator for Energy-Efficient Edge Vision Applications. *Journal of Reconfigurable Hardware Architectures and Embedded Systems*, 3(3), 9–16.
13. Pires, T., Schlinger, E., & Garrette, D. (2019, July). How multilingual is multilingual BERT? In *Proceedings of the 57th annual meeting of the association for computational linguistics* (pp. 4996-5001). <https://doi.org/10.18653/v1/P19-1493>
14. Raganato, A., Camacho-Collados, J., & Navigli, R. (2017, April). Word sense disambiguation: A unified evaluation framework and empirical comparison. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers* (pp. 99-110).
15. Raganato, A., Pasini, T., Camacho-Collados, J., & Pilehvar, M. T. (2020, November). XL-WiC: A multilingual benchmark for evaluating semantic contextualization. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 7193-7206). <https://doi.org/10.18653/v1/2020.emnlp-main.584>
16. Rustamov, D., Safarov, D., Mustafojeva, N., Xaitov, Z., Khabibullina, L., Narzullaeva, F., & Mamatova, F. (2026). Neuroscientific Perspectives on Multilingual Philology and Cognitive Development in Academic Settings. *Indian Journal of Information Sources and Services*, 16(2), 180–187. <https://doi.org/10.51983/ijiss-2026.16.2.19>
17. Mukti, I. Z., Khan, E. R., & Biswas, K. K. (2026). Attack-Resilient Quantized Secure Control for Nonlinear Distributed Automation Systems. *Transactions on Internet Security, Cloud Services, and Distributed Applications*, 31-38.