Typological Similarity and Fine-Tuning Strategies in Cross-Lingual Euphemism Detection

Assignee Research

June 29, 2026

Abstract

Euphemisms are culturally variable and often ambiguous, posing challenges for language models, especially in low-resource settings. This paper investigates how cross-lingual transfer via sequential fine-tuning affects euphemism detection across five languages: English, Spanish, Chinese, Turkish, and Yoruba. We compare sequential fine-tuning with monolingual and simultaneous fine-tuning using XLM-R and mBERT, analyzing how performance is shaped by language pairings, typological features, and pretraining coverage. Results show that sequential fine-tuning with a high-resource L1 improves L2 perfo

1 Introduction

This paper examines: Sequential Fine-Tuning with XLM-R for Cross-Lingual Euphemism Detection Across Varying Typological Distances. Research question: Does the typological similarity between source and target languages mitigate the performance gap in cross-lingual euphemism detection when using simultaneous vs. sequential fine-tuning in models like XLM-R or mBERT?.

2 Methodology

Systematic literature search across multiple databases yielded 5 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.8/10.

3 Results

5 papers retrieved. 8 claims extracted; 8 independently verified. Quality review score: 8.8/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
Euphemisms are culturally variable and often ambiguous, posing challenges for language models, especially in low-resourc	✓	0.34
This paper investigates how cross-lingual transfer via sequential fine-tuning affects euphemism detection across five la	✓	0.47
Sequential fine-tuning with a high-resource L1 improves L2 performance in euphemism detection.	✓	0.35
The research goal is to measure the impact of typological similarity between source and target languages on cross-lingua	✓	0.49
The study compares sequential fine-tuning with monolingual and simultaneous fine-tuning using XLM-R and mBERT.	✓	0.32
Performance in euphemism detection is shaped by language pairings, typological features, and pretraining coverage.	✓	0.28
The study analyzes accuracy differences between language pairs with varying typological distances.	✓	0.28
The autonomous synthesis report was generated by Assignee Research with a tribunal consensus score of 8.3/10.	✓	0.33

References

- https://doi.org/10.26615/978-954-452-098-4-122
- https://doi.org/10.5281/zenodo.20843631
- https://doi.org/10.5281/zenodo.20843632