CURRENT RESEARCH JOURNAL OF PHILOLOGICAL SCIENCES
(ISSN –2767-3758)
VOLUME 04 ISSUE 08 Pages: 6-10
SJIF IMPACT FACTOR (2021: 5. 823) (2022: 6. 041) (2023: 7. 491)
OCLC – 1242423883
Publisher: Master Journals

Journal Website: https://masterjournals.com/index.php/crjps

Research Article

ADDRESSING NOISE ISSUES IN SENTENCE STRUCTURE FOR MORPHOLOGICAL ANALYSIS OF ENGLISH LANGUAGE SENTENCES FOR HINDI LANGUAGE USERS

Submission Date: July 28, 2023, **Accepted Date:** Aug 02, 2023,
Published Date: Aug 07, 2023
Crossref doi: https://doi.org/10.37547/philological-crjps-04-08-02

Richa Mehta
Professor of Language Teaching and Hindi Literature at Lucknow University, India

ABSTRACT

Morphological analysis is a fundamental task in natural language processing that involves breaking down words into their constituent morphemes to understand their grammatical structure and meaning. However, when morphological analysis is applied to English language sentences by Hindi language users, noise issues arise due to the syntactic and structural differences between the two languages. This study addresses these noise issues and proposes techniques to improve the accuracy of morphological analysis for English sentences processed by Hindi language users. By exploring methods to handle word order variations, phrasal differences, and other syntactic disparities, this research aims to enhance the effectiveness of morphological analysis tools for bilingual users and facilitate their comprehension of English language sentences.

KEYWORDS

Morphological analysis, noise issues, sentence structure, English language sentences, Hindi language users, natural language processing, bilingual users, word order variations, syntactic disparities, phrasal differences, grammatical structure, meaning, linguistic differences.

INTRODUCTION

Morphological analysis is a crucial component of natural language processing, enabling the understanding of word structure and grammar within sentences. For bilingual users, such as Hindi

CURRENT RESEARCH JOURNAL OF PHILOLOGICAL SCIENCES
(ISSN –2767-3758)
VOLUME 04 ISSUE 08 Pages: 6-10
SJIF IMPACT FACTOR (2021: 5. 823) (2022: 6. 041) (2023: 7. 491)
OCLC – 1242423883
Publisher: Master Journals

language speakers analyzing English language sentences, morphological analysis becomes challenging due to the differences in sentence structures and syntactic rules between the two languages. These structural disparities introduce noise issues, leading to inaccurate morphological analysis results and hindering the comprehension of English sentences by Hindi language users.

This research aims to address the noise issues in sentence structure for morphological analysis of English language sentences for Hindi language users. By identifying and mitigating the challenges arising from linguistic differences, this study seeks to enhance the accuracy and effectiveness of morphological analysis tools for bilingual users. The insights gained from this research can benefit various natural language processing applications, including machine translation, information retrieval, and language learning tools.

METHOD

Data Collection:
A corpus of English language sentences is compiled from various sources, and a subset of these sentences is annotated for morphological analysis. Hindi language users are consulted to validate the selection of sentences that present typical noise issues encountered during morphological analysis.

Linguistic Analysis:
The linguistic disparities between English and Hindi sentence structures are identified and analyzed. Key differences related to word order, phrasal construction, verb-noun agreement, and other syntactic rules are documented.

Noise Issue Identification:
The annotated subset of English sentences is analyzed to identify specific noise issues that arise when Hindi language users attempt morphological analysis. Common challenges include misinterpretation of verb forms, incorrect word segmentation, and difficulties in recognizing compound words and idiomatic expressions.

Noise Mitigation Techniques:
Various techniques are explored to address the identified noise issues in morphological analysis. These techniques include language-specific rule-based algorithms, statistical models, and machine learning approaches trained on bilingual data.

Evaluation:
The effectiveness of the proposed noise mitigation techniques is evaluated on a test set of English language sentences processed by Hindi language users. The evaluation metrics include accuracy, precision, recall, and F1 score to measure the improvement in morphological analysis accuracy.

User Feedback:
Hindi language users participate in usability testing to provide feedback on the effectiveness and user-friendliness of the proposed noise mitigation techniques. User feedback helps refine

CURRENT RESEARCH JOURNAL OF PHILOLOGICAL SCIENCES
(ISSN –2767-3758)
VOLUME 04 ISSUE 08 Pages: 6-10
SJIF IMPACT FACTOR (2021: 5. 823) (2022: 6. 041) (2023: 7. 491)
OCLC – 1242423883
Publisher: Master Journals

and optimize the techniques based on practical application and user needs.

Comparative Analysis:

The performance of the proposed noise mitigation techniques is compared with existing morphological analysis tools designed for bilingual users. The comparative analysis aims to demonstrate the advantages and novel contributions of the proposed approach.

By employing a systematic method that combines linguistic analysis, noise issue identification, and evaluation with user feedback, this research seeks to advance the field of morphological analysis for bilingual users. The proposed techniques aim to enhance the accuracy and efficiency of morphological analysis for English language sentences processed by Hindi language users, ultimately facilitating their comprehension and interaction with English text.

RESULTS

The evaluation of noise mitigation techniques for morphological analysis of English language sentences processed by Hindi language users showed promising results. Several noise issues commonly encountered by Hindi language users during morphological analysis were addressed effectively. The proposed techniques demonstrated improved accuracy in capturing the underlying morphological structure of English words and sentences, reducing the impact of syntactic disparities between English and Hindi.

DISCUSSION

The linguistic analysis revealed significant differences between English and Hindi sentence structures, leading to challenges for Hindi language users in accurately analyzing English language sentences. Common noise issues included difficulties in recognizing verb forms, handling compound words, and dealing with word order variations. These challenges often resulted in incorrect morphological analysis outcomes and impeded the comprehension of English text.

The noise mitigation techniques were designed to tackle these specific challenges. Rule-based algorithms were employed to address word segmentation and verb form recognition. Statistical models and machine learning approaches were applied to improve the handling of compound words and phrasal constructions. The combination of these techniques effectively reduced the impact of linguistic disparities, resulting in more accurate and contextually relevant morphological analysis.

User feedback provided valuable insights into the usability and effectiveness of the proposed techniques. Hindi language users appreciated the improvements in morphological analysis accuracy, as it facilitated their understanding of English text. The user feedback also highlighted the need for user-friendly interfaces and interactive features to support bilingual users effectively.

CONCLUSION

CURRENT RESEARCH JOURNAL OF PHILOLOGICAL SCIENCES
(ISSN –2767-3758)
VOLUME 04 ISSUE 08 Pages: 6-10
SJIF IMPACT FACTOR (2021: 5. 823) (2022: 6. 041) (2023: 7. 491)
OCLC – 1242423883
Publisher: Master Journals

This research successfully addressed noise issues in sentence structure for morphological analysis of English language sentences for Hindi language users. By employing a systematic approach that included linguistic analysis, noise issue identification, evaluation, and user feedback, the proposed techniques demonstrated significant improvements in morphological analysis accuracy.

The findings of this study contribute to the development of more effective natural language processing tools for bilingual users. By acknowledging and mitigating the linguistic disparities between languages, morphological analysis tools can better cater to the specific needs of Hindi language users analyzing English text.

The proposed noise mitigation techniques hold promise for enhancing various natural language processing applications, including machine translation, information retrieval, and language learning tools. Furthermore, the research highlights the importance of considering the linguistic characteristics and user perspectives when designing tools for bilingual users.

In conclusion, the study provides practical solutions to address noise issues in morphological analysis for Hindi language users dealing with English language sentences. The proposed techniques offer valuable insights for researchers and developers working on multilingual natural language processing tasks, ultimately contributing to the advancement of language technology and facilitating cross-lingual communication and comprehension.

REFERENCES

1. R. Mahesh, K. Sinha, and A. Thaku, "How to get best results out of a machine translation system: A case study of English to Hindi Translation," CSI Journal, vol. 38, no. 4, Oct.-Dec. 2008.
2. L. V. Subramaniam, S. Roy, T. A. Faruquie, and S. Negi, "A survey of types of text noise and techniques to handle noisy text," in Proc. The Third Workshop on Analytics for Noisy Unstructured Text Data, ACM, 2009, pp. 115-122.
3. TDIL. Machine translation. Indian Language Technology Proliferation and Deployment Center. [Online]. Available: http://tdil-dc.in/components/com_mtsystem/CommonUI/homeMT.php
4. H. Trost, X2MORF: A morphological component based on augmented two-level morphology, Research Report, 1991.
5. James, Natural Language Understanding, 2nd ed. Pearson Education, 2008.
6. H. Schmid, "A Programming language for finite state transducers," in Proc. the 5th International Workshop on Finite State Methods in Natural Language Processing, Helsinki, Finland, July 13, 2005, pp. 308-314.
7. H. Schmid. Developing computational morphologies with the SFST tools. Tutorial SFST Tool. [Online]. Available: http://www.cis.uni-

CURRENT RESEARCH JOURNAL OF PHILOLOGICAL SCIENCES
(ISSN –2767-3758)
VOLUME 04 ISSUE 08 Pages: 6-10
SJIF IMPACT FACTOR (2021: 5. 823) (2022: 6. 041) (2023: 7. 491)
OCLC – 1242423883
Publisher: Master Journals

muenchen.de/~schmid/tools/SFST/data/SFST-Tutorial.pdf

8. R. M. K. Sinha and A. Jain, "Angla Hindi: an English to Hindi machine-aided translation system," MT Summit IX, New Orleans, USA, pp. 494-497, 2003.
9. D. Jurafsky and J. H. Martin, Speech and Language Processing, 2nd ed. Prentice Hall Inc., 2002.