# **Facilitating Language Hacking with Digital Tools: A Study of Translation Alignment**

A thesis  
submitted by

Joseph C. Hilleary

In partial fulfillment of the requirements  
for the degree of

Master of Science  
in  
*Computer Science*

TUFTS UNIVERSITY

May 2024

© 2024, Joseph C. Hilleary

Advisor: Gregory Crane

# Abstract

This project examines digital translation alignment as a tool for facilitating a new paradigm of “language hacking.” In this frame, readers use translations as an aid for engaging directly with source texts in languages with which they are unfamiliar, rather than as a replacement for the original materials. I present the results of a pilot study measuring the impact of this form of annotation on the comprehension of users of a digital reading environment and introduce and evaluate a novel approach to the task of generating aligned translations using Universal Dependency Trees.

# Acknowledgements

# Table of Contents

# List of Tables

# List of Figures

# Introduction

The emerging field of “language hacking” seeks to establish an intermediate paradigm in the humanities between working only with texts in languages over which one has complete mastery and exclusively reading materials in translation. In particular, recent work has promoted the proliferation of word-level aligned translations as a tool for exposing readers directly to source materials, while providing them with support to make sense of languages with which they may be unfamiliar or uncomfortable.

The use of annotation and translation alignment to facilitate the understanding of texts is not new. For centuries, humanist scholars have created critical editions rife with marginalia and notations to assist readers in gaining a deeper appreciation for the nuances of the material. What marks a profound shift in the field, however, is the advancement of digital technologies enabling the rapid scaling of such annotations. In the past, the production of translations, alignments, and other forms of annotation was extremely time consuming. Today, computational tools, such as machine learning, allow some of this work to be done on the fly.

Furthermore, a contemporary language hacker does not need to be a scholar looking to delve into a historic canonical text. In fact, a more common example might be that of a native speaker of a regional language attempting to navigate a globalized world through machine translation. Tools like Google Translate, DeepL, and others that rely on new neural architectures have made huge strides in accuracy over the last few years, but remain rife with errors, especially for lower resource languages. Sometimes these translation errors are obvious, but without recourse to the original text, they can be difficult for readers to work around. Even when such translations are technically correct, they, like all translations, lack some of the nuance and semantic aspects present in the original text.

In both framings—that of the philologist and the machine translation-reliant citizen—the key is to provide tools that enable the reader to engage with the original text while using the translation as an aid rather than a replacement. This project seeks to explore the use of digital translation alignment as a tool for language hackers of all types. In the first section, building on work looking at the value of translation alignment as a pedagogical tool, I present a study providing empirical evidence of the impact of word-level alignment on reading comprehension in a language unknown to the reader. Having established a justification for presenting readers with word-level alignments of translations, in the second section I propose a novel approach to the task of automatically generating word-aligned translations using Universal Dependency Trees.

# Chapter 1: Impact of Word-Level Aligned Translations on Reader Comprehension of an Unknown Language

## Related Work

### Use of Alignment for Language Learning

### Use of Cloze Tests for Evaluating L2 Comprehension

## Experimental Design

## Results

## Conclusion

# Chapter 2: A Novel Approach to Translation Alignment Using Universal Dependency Tree Matching

## Related Work

### Automated Translation Alignment

### Universal Dependencies

## Algorithm Design

## Experiments

## Results

## Conclusion

# Conclusion

# Appendix

# Endnotes

# Bibliography