

# LANGUAGES WE CAN'T READ

Advanced Localization Project Management  
Final Report

## OUR MISSION

We provide information for marginalized and minority groups, including those who speak endangered languages, for we believe all locales deserve access to information--no matter near or far. We strive to provide accurate translations with speedy turnaround, especially in case of global emergencies and humanitarian crises.

## Team Penguins

Andrew Taylor, Elena Malamed, Luke Paustian,  
Olivia Plowman, Ruby Lee

## Table of Contents

Introduction .....	2
Critical Analysis .....	2
Research .....	4
Best Practices.....	5
Conduct Preliminary Research:.....	5
Pay Attention to Punctuation: .....	5
Look for Error Characters:.....	6
You Can Ask for Help: .....	6
Conclusion.....	6
Works Cited .....	6

## Introduction

The process of managing a localization project is a detailed-oriented work, in which one has to look out for any mistakes, big or small, in order to deliver the final project at its best quality. Now, what if we can't read or understand the language we're managing? How can we make sure all the possible details are taken care of? Managing an unfamiliar language is one of the key takeaways our team acquired for this project. Our team has been working on the localization of Milwaukee Public School's website throughout the semester. Since languages serve as a critical element in terms of acquiring information, it is only fair if people all have the same opportunity to access content that they can understand. With this in mind, we evaluated the location and the cultural background of the Milwaukee Public School, and decided to localize the school's website into 4 languages: Burmese, Karen, Simplified Chinese, and Spanish, and focused especially on Burmese and Simplified Chinese. Among these languages, none of them were our first or second language.

## Critical Analysis

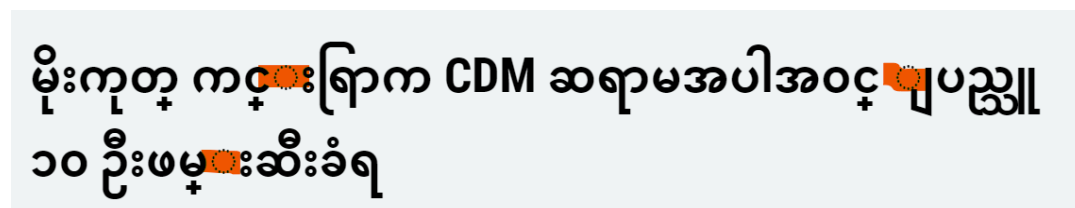
One of the most difficult aspects of our project was the challenge of post-editing and reviewing content in exotic alphabets that we knew little to nothing about. How could we be sure that the segments were fully translated? That tags and formatting was not altered during machine translation? Or that terminology was consistent throughout?

While it is unrealistic to expect perfect translations from a post-editor who does not speak the language, there are some steps that one can take to eliminate clearly visible errors. In our experiences with translating into Burmese, one of the issues that we found was that some words in the Memsources window were strangely indented, creating large empty spaces when there was no indentation in the source language. Some quick research and referencing a separate machine translation engine taught our group that Burmese requires no spaces between words, so we did our best to eliminate the unnecessary spaces and ensure that the preview on html looked appropriate.

Some other issues that can occur include improper segmentation, such as a

sentence being cut into two separate segments due to an abbreviation containing a period such as “Mr.”, or having certain words that are left untranslated in the target language due to there not being an appropriate translation, such as the names of U.S. states. We found that some of these issues can be easily settled, such as recognizing a broken sentence from the source and splicing the two parts together. However, when issues require knowledge of the language to resolve, one must either do their research or simply ask a speaker of the language.

Researching the basics of the writing system, alphabets, formatting customs, etc. of a language can vastly illuminate the most obvious problems. By taking time to look at Burmese articles and websites we became slightly more accustomed to how the language should look. For example, in our Burmese translation, some of our group members noticed that their Burmese font was having Unicode errors. This caused for some part of the characters to be replaced with dotted circles:



By comparing our translated text with native Burmese text that we found online, these Unicode problem markers became red flags for further issues in our project. Not only being familiar with the language, but being familiar with how text can be corrupted is also useful for post editing in cases where the language is not familiar.

We also were able to tackle our post editing by focusing on what we knew and using context clues from the source text. For example, we may not have been able to read the Burmese text, but where English text remained untranslated, or Romanized we were able to gain context clues about the words around it. In the example below, there were many cases where certain words were repeatedly left untranslated, or were translated inconsistently.

543	1 2 3 4 Seal of Biliteracy Schools 5 2	4 5 Seal of Biliteracy	1 2 3 5 Biliteracy ကျောင်းများ၏တံဆိပ် 5 4 ခတ် Biliteracy ကျောင်းများ၏တံဆိပ် 4 2	✓
544	1		1	✓
545	1 Seal of Biliteracy 1		1 Biliteracy ၏တံဆိပ်ခတ် 1	✓
546	MPS students can now earn the Seal of Biliteracy!		MPS ကျောင်းသားများသည်ယခုအခါ Biliteracy ၏ Seal ကိုရရှိနိုင်သည်။	✓

The phrase “Seal of Biliteracy” appeared in multiple different segments, and by comparing the words that were left untranslated, inconsistencies became clearer. In Segment 543 “seal” is translated but “biliteracy” is not. This is the same for segment 245. However, in segment 546 both words remain untranslated. By reviewing the source text, it becomes clear that “Seal of Biliteracy” between these three segments all should have been translated uniformly. These types of error can then be flagged and sent back to the translator for review.

## Research

But what do other people have to say about editing or managing a project in a foreign language? We aren’t the first people who have had to do so, and we most certainly won’t be the last.

Casimir Nozkowski, although a film editor and not a localization project manager or translator, was interviewed by postPerspective about his work on an animated documentary. The short film, called *The Shining Star of Losers Everywhere*, was recorded with Japanese audio, and Nozkowski had never “cut a film in a language other than English”. In the interview, Nozkowski recalls how much the translators contributed to the project through transcription software and translation that resulted in subtitles being put in “seamlessly” at the right times. On the other hand, he also mentions about how he often flubbed the editing process:

I’m so used to working in English and being able to work a little editing magic on interviews where you’re not changing what someone’s saying but you’re able to kind of speed them up or help them say something more efficiently. With our subjects speaking Japanese, and the sentence structures being a bit different, I was often just guessing on which words I did or did not need to make a succinct point. I was often wrong. Doh! But luckily our translators were really with us throughout the process and kept us in the clear.

Nozkowski, as a post-production editor and not someone trained to work with another language (at least, not at the time, as this interview was conducted in 2016), had to rely a lot on the translators on the project. However, that in itself is a good lesson for localization project managers—when in doubt, don't be afraid to ask for help.

In an interview with No Film School, post-production editors Na'Imah Pope and Sara Sachs spoke about their experiences editing audio/visual projects in foreign languages. The two both had experience editing in foreign languages multiple times, they explained, so they were not completely fresh to the difficulty. Other than taking advantage of syncing and transcription software, one of them—Pope—urged the importance of organization. “As an editor and assistant editor, being organized is tremendously important even if you speak the language, but more so when you do not,” she added. They also recommended to really “study” the project while working on it to notice patterns of speech.

## Best Practices

As a wrap up to our experiences and findings with managing a localization project in a foreign language we could not understand, we thought it would be good to create a short list of best practices for anyone who finds themselves working with a language they do not understand in the future.

### Conduct Preliminary Research:

Finding a basic guide describing the grammatical structures, alphabet, and proper punctuation is easy to do via the internet, even for many remote and marginalized languages. Find these web pages, and keep their locations saved on your internet browser, or on your device as an HTML file. Try your best to make sure that your sources are trustworthy and consider sharing these sources with a trusted speaker of the language in question to confirm their quality as a reference.

### Pay Attention to Punctuation:

Even in a language you are not fluent in, you should be able to use punctuation marks to get a general idea of sentence structure. If any punctuation stands out as odd, take note of it. Use punctuation to identify individual sentences and compare them to the source/target text. Although not technically punctuation

marks, if you are working with a text that contains HTML tags, know that these can also be used to help identify specific phrases and terms in the text.

#### Look for Error Characters:

Everyone knows what traditional tofu boxes look like, but also look for non-standard mojibake including dotted circle characters ◦ and question blocks ❖.

Similarly, if you see a word using characters from a different language, take note. If its an untranslated word, try to find the official translation.

#### You Can Ask for Help:

If you know someone, or your office has a contract with someone, who knows anything about the language you are working in, do not be afraid to reach out. You can go a long way to understanding how to identify potential problems with research and punctuation tricks, but if in doubt, always reach out, especially if perfection is an expectation of the project.

## Conclusion

It can seem very intimidating, potentially even impossible, to manage and edit in a language that is not your native tongue. At the end of the day, however, this is something almost any project manager will deal with on a regular basis. In our experience, we found that being able to identify patterns and conduct swift research were essential skills in working with a foreign language. Of course, it can be difficult to tell if you are even doing well when editing on this type of job; however, we are happy to say that we believe the experience of this project has taught us the skills we need to prepare for working with a language we cannot read, and the confidence to face this kind of project if we encounter something like it again in our future careers.

## Works Cited

Kleinman, Lily. "Invaluable Tips for Editing in a Foreign Language." *No Film School*, No Film School, 1 Mar. 2021, [nofilmschool.com/invaluable-tips-editing-foreign-language](https://nofilmschool.com/invaluable-tips-editing-foreign-language).

MustGo. "Burmese Language - Structure, Writing & Alphabet - MustGo."

*MustGo.com*, [www.mustgo.com/worldlanguages/burmese/](http://www.mustgo.com/worldlanguages/burmese/).

Pregot, Kristine. "Casimir Nozkowski: The Challenges of Editing a Foreign-Language

Doc." *PostPerspective*, Randi Altman, 26 May 2016,

[postperspective.com/shining-star-the-challenges-of-editing-in-another-language/](http://postperspective.com/shining-star-the-challenges-of-editing-in-another-language/).