Eleni Vligouridou
Joaquin Cherta Rodriguez

Project Report - Meaning Representations

1.   Introduction

The purpose of this project was to investigate AMR structures in three different languages. By creating AMR annotations in Catalan and Greek from The Little Prince corpus and comparing them with the already published English AMR annotations, we aim to demonstrate how similar these languages are in terms of semantics, using the Smatch score for comparison.

2.   Time tracker

Task	Hours
Regular meetings	9
First annotations	25
Literature review and discussions	12
Alignment	4
PropBank mapping and first annotation review	8
Annotation mapping to English	4
Challenges with smatch parsing	10
Final annotation review, results and report	8
Total hours	**80**

3.   Stages of the project

3.1 Regular meetings

During the whole duration of the project, we had regular meetings for planification, to clear possible doubts and to begin every stage of the project from the same standpoint.

3.2 First annotations

We performed our first hands-on work with the data, and we originally annotated 50 sentences extracted from the book of *The Little Prince*, written by Antoine de Saint-Exupéry. Eleni annotated the Greek version, and Quim annotated the Catalan version. We used an already annotated version for English. As a result, we had three .txt files containing the AMR annotated sentences with the Penman notation.

3.3 Literature review and discussions

We weren't sure how to proceed with our project, so we checked different papers. We struggled to find a clear path or working strategy to further advance. We strongly considered working with approaches like Translate then parse, or DeAMR, but realised after more discussion and consideration we didn't have the tools or time to expand the T+P work with our own languages, and likewise we found it was too ambitious to be able to create a set of guidelines that would create our versions for ElAMR and CatAMR. Finally we decided to compare our annotations structurally by mapping them to English while keeping the original Catalan and Greek structure, so that we could properly parse them.

3.4 Alignment

We needed to align the sentences of the Greek, Catalan, and English versions because the different translations led to different structures for each version, so we had to make sure that the annotations were comparable and corresponded from one language to another. For that, we had to modify all of the three versions, and we did it with the following criteria:

- When we encounter an alignment problem, we always modify the least amount possible of annotations (if two languages are matching, we modify the third one which doesn't match).

- When sentences are not aligned due to coordination or punctuation, we split or bond sentences together. We punctuate and add caps to the split sentences.

- Two sentences were also manually added to the Catalan version (chapter headings) because they weren't present (stylistic differences in the book, where in Catalan the chapter was introduced with a drawing, and in Greek and English with text).

- We skipped one sentence in the English and Catalan data, since it was not present at all in the Greek version. We decided that adding it in Greek would create too much of an impact/interference from the annotators on the original text, in contrast with the issue with headings in Catalan, which were much easier to handle.

3.5 PropBank mapping and first annotation review

We used PropBank to map the Catalan and Greek verbs (and verb senses) to English ones. There was the option of using AnCora to assign senses to Catalan verbs, but we decided against it to keep the consistency between the three languages since Greek does not have a resource like that, even if that implied that the verbs and senses from PropBank were harder to adapt to Catalan.

Matching our annotations to the PropBank verb structures involved correcting and changing many annotations. More than half of our annotations were partially changed, and some were completely remade. We also had to newly annotate all the sentences that were changed because of the alignment process. Finally, we made a formatting revision to ensure that our annotations were correct and well formatted.

During this process, we encountered some linguistic difficulties produced by the differences between our languages:

- Ellipsis: Catalan and Greek are pro-drop languages, so we needed to make them explicit in our annotation, even if they were not in the sentence. Depending on the context, we annotated them by using a pronoun or the noun referent from previous sentences.

- Phrasal verbs: we had trouble annotating English phrasal verbs, since most of the time the structure and rolesets from PropBank couldn't be replicated in our Catalan and Greek annotations. Instead, we tried to find a similar verb that would better match the original structure while keeping a similar meaning.

3.6 Annotation mapping to English

We used ChatGPT to speed up the annotation translation process. We asked ChatGPT to do a preliminary translation for us sentence by sentence, which we reviewed and modified afterwards. We experienced some problems with the use of ChatGPT, so these translations were also manually done. This step also required another formatting review, since the process created format changes and problems, like badly placed parentheses or changed variable names.

3.7 Challenges with smatch parsing

Before parsing the sentences, we did a final format review by using a script to review the formatting of Penman notation, and we fixed some mistakes. We had never worked with Smatch before, so we reviewed the documentation and investigated how to use it.

After that we encountered some difficulties and problems with our parsing. We solved some encoding problems when processing our files, but we were still not able to parse our annotations. For that, we created test files with smaller subsamples of our annotations. We spent some time researching all the possibilities to solve our problem, and finally we isolated it and found that there were still some formatting problems that the previous check with the script didn't solve, so we went back to the annotations to review them for the last time.

3.8 Final annotation review, results and report

After we understood the types of errors we were encountering during Smatch, we had to change our annotations one last time. There were some duplicated variables in the file and some mismatched parentheses. After we resolved all of these issues, we were able to run Smatch and get the F1 scores. In the end, we put our scores into a table and interpreted the results in the report. We wanted to present our work in a PDF file, so we created a document where we briefly described our work, the difficulties we faced, and our interpretation of the results.

4.   Results and conclusion

As mentioned earlier, in order to examine the semantic differences between Catalan, Greek, and English, we decided to use a Python library that computes the Smatch

score, which measures how similar two AMR graphs are. Using the English AMR annotations as the gold standard and the Catalan and Greek AMR annotations as the test data, we got the following scores:

Test	Gold	F1-score
Catalan	English	0.63
Catalan	Greek	0.60
Greek	English	0.62

As you can see, the scores range from 0.60 to 0.63. The highest similarity is shown between Catalan and English, while the lowest similarity is between Catalan and Greek, despite the fact that we agreed on many aspects during the annotation process. The similarity between Greek and English is close to that of the Catalan-English comparison.

We interpret the results as moderate similarity. There are core elements that are shared, the meaning is generally aligned, but the representations still diverge. This is something we also observed during the annotation process. In general, we would conclude that the AMR graphs are aligned, but there are semantic differences and varying interpretations, maybe also due to translation.