

# EVALUATING MACHINE TRANSLATION OF SUBJUNCTIVE

Behnoosh Namdarzadeh (behnooshnamdar@gmail.com)

Steven Saily (steven.saily@gmail.com)

## Abstract

Our objective is to evaluate the translation of subjunctive verbs from French to English and from English to French, especially in concessive or volitive constructions, and after prepositions.

In English, verbs do not have a specific form for the subjunctive mood but use the same construction as imperative or infinitive moods. Thus, a machine has to detect when subjunctive must be used, even if the verb in English is written as if it was using the infinitive. Translation from French to English faces the same problem: verbs at subjunctive mood in French may use the same construction as in the infinitive.

## Procedure

### 1 Corpus selection

The *challenge set* is manually constructed and consists of 116 sentences in French and 114 sentences in English, extracted from [?], [?], [?], [?], [?] and [?]. We chose them because they could be problematic when translated, as they contain (in French) or are supposed to contain (when translated from English to French) structures such as “bien que” or “préférer que”. The *target set* consists of the *challenge set* manually translated. It will be the corpus used to evaluate the automatic translations of the sentences from the *challenge set*.

### 2 Evaluation method

To evaluate the automatic translations, we will tag parts supposed to be subjunctive in both target sentence and translated sentence. Then, we will compare both tags and sentences meanings, and a score  $\delta$  will be given to the translation.

Translation	$\delta$
Correctly used subjunctive (C)	+1
Correct sentence but without subjunctive (CW)	+0.5
Correct subjunctive but incorrect sentence (CI)	-0.5
Incorrect (I)	-1

The *score of a translated set*  $\Delta$  will be defined as the sum of the scores given to the sentences translated by this model:

$$\Delta_S = \sum_{s \in S} \delta_s$$

( $s$  a sentence,  $S$  the translated set). Thus:  $\forall S, |\Delta_{S(\text{French} \rightarrow \text{English})}| \leq 116$ ,  $|\Delta_{S(\text{English} \rightarrow \text{French})}| \leq 114$ . The *correctness of a translated set*  $\tau$  will be defined as the rate of correctly translated sentences:

$$0 \leq \tau_S = \frac{|C_S| + |CW_S|}{|S|} \leq 1$$

### 3 First translations and evaluations

The first translations will be made using Google Translate graphic interfaces. We will then evaluate those translations. This set of translations will be referred as the *GUI set*.

### 4 Translating using mBART

A new translated set, referred as the *mBART set* will be obtained using the pretrained *mBART-50 many to many multilingual machine translation* model. [?] It will be compared to both the *target set* and the *GUI set*.

## Introduction

Neural Machine Translation (NMT) outperforms other systems tackling with long sentences [?] and have the capacity to learn idiosyncratic linguistic phenomena and train themselves.[?] Yet, deep neural models are embedded with tens of thousands of neurons and millions of parameters which make it hard to predict its performance in action. Hence, the research in the field is open to replication to better understand the algorithms within.[?] In the current study, we intend to test the system encountering one of the challenging moods in both French and English, subjunctive.

## Theoretical Background

The pioneering study of MT systems might be the study of Sennrich [?], who built challenge sets focusing on five language characteristics. In the wake of Sennrich (2017), Isabelle [?] [?] proposed his challenge set model, in which he introduced four levels of language difficulties embedded in translation at morpho-syntactic, lexico-syntactic, syntactic, and lexical levels. As mentioned earlier, subjunctive mood is the main focus of this topic which is categorized at the syntactic level.

## Results

French to English sets:

$$\Delta_{GUI \text{ set}(\text{French} \rightarrow \text{English})} = 46.5 ; \frac{\Delta_{GUI \text{ set}(\text{French} \rightarrow \text{English})}}{\max(\Delta_{S(\text{French} \rightarrow \text{English})})} = 0.41 ; \tau_{GUI \text{ set}(\text{French} \rightarrow \text{English})} = 0.78$$

$$\Delta_{mBART \text{ set}(\text{French} \rightarrow \text{English})} = 51.5 ; \frac{\Delta_{mBART \text{ set}(\text{French} \rightarrow \text{English})}}{\max(\Delta_{S(\text{French} \rightarrow \text{English})})} = 0.45 ; \tau_{mBART \text{ set}(\text{French} \rightarrow \text{English})} = 0.82$$

English to French sets:

$$\Delta_{GUI \text{ set}(\text{English} \rightarrow \text{French})} = 73.5 ; \frac{\Delta_{GUI \text{ set}(\text{English} \rightarrow \text{French})}}{\max(\Delta_{S(\text{English} \rightarrow \text{French})})} = 0.63 ; \tau_{GUI \text{ set}(\text{English} \rightarrow \text{French})} = 0.88$$

$$\Delta_{mBART \text{ set}(\text{English} \rightarrow \text{French})} = 66 ; \frac{\Delta_{mBART \text{ set}(\text{English} \rightarrow \text{French})}}{\max(\Delta_{S(\text{English} \rightarrow \text{French})})} = 0.57 ; \tau_{mBART \text{ set}(\text{English} \rightarrow \text{French})} = 0.83$$

## Conclusion

Though the different  $\tau$  scores are satisfying, the  $\Delta_{S(\text{French} \rightarrow \text{English})}$  scores are disappointing as they do not reach half of their possible maximums, while the  $\Delta_{S(\text{English} \rightarrow \text{French})}$  scores are by far more satisfying. The differences between  $\Delta$  and  $\tau$  scores are quite high. This can be explained by the utilisation of correct formulation which do not require the subjunctive mood. Among them, we can cite *falloir que* replaced by *devoir + infinitive form* (see Table 2). We can also notice that some structures which may or may not require a subjunctive depending on the sentence are well-handled by both translation systems.

## Bibliography



This project was supported by IdEx Université de Paris ANR-18-IDEX-0001 (AAP Initiation à la Recherche en Licence, vague 2)