

# Guidelines for annotating machine translation errors using the SCATE error taxonomy

To be able to assess the quality of Neural Machine Translation systems all errors in the MT output will be indicated and labelled.

To be able to judge the quality of the translations, they will be marked for two important aspects: **accuracy and fluency**. **Accuracy** is concerned with the relationship between source text and target text, whereas **fluency** is concerned with the target text and language.

You will use a two-step approach. In a first step only the translation (target sentence) will be visible and all **fluency errors** will be annotated; in a second step, the source sentence will be displayed as well and you have to compare the source and target sentence and indicate all accuracy errors. There can be overlapping fluency and accuracy errors (also within one of the two categories too). To label the errors, you will make use of a predefined error taxonomy. Please annotate errors by annotating minimal text spans, in which the error itself is visible (but the span does not contain other tokens, which are not a part of the error).

## Step 1: Fluency errors

Fluency						
Coherence	Lexicon	Grammar & Syntax	Style & Register	Orthography	Multiple Errors	Other
Logical problem	Lexical choice	Agreement	Disfluent sentence or construction	Capitalization		
Discourse marker	Wrong preposition	Verb form	Repetition	Punctuation		
Co-reference	Non-existing/foreign word	Word order	Register	Spelling		
Inconsistency		Extra words		Other		
Verb tense		Missing words				
		Other				

The goal of the first step is to annotate the machine translated text for fluency. Fluency can be described as respecting the norms of the target language and culture. A good translation should read as a native Dutch text. This includes respecting the conventions of the language (grammar, lexicon, spelling) as well as respecting the conventions of the text structure (paragraph content and coherence) and the text type, in this case literature.

As an annotator, it is your task to mark anything that does not follow the conventions of the Dutch language or that does not respect the demands of the text structure or the text type. To facilitate the task, different categories (Coherence, Lexicon, Style & register, Grammar & Syntax, Orthography and Multiple errors) are predefined in WebAnno, along with their most important subcategories.

## Coherence

There is something wrong with the coherence of the text: lack of logical structure, confusing relationships, ...

Logical problem: the information as such is illogical/confusing/makes no sense when looking at the rest of the text

Logical problem | Coherence  
We hadden een goed garen over oude tijden en het eindigde met zijn uitnodiging om naar Styles te gaan om daar mijn verlof door te brengen.

Logical problem | Coherence  
"De mater zal blij zijn om je weer te zien ... na al die jaren," voegde hij eraan toe.

Logical problem | Coherence  
"Ik heb een neef die borstvoeding geeft," merkte ik op.

Cultural reference: the word/phrase refers to something specific of the culture of the source, but makes no sense in the target language.

Cultural reference | Coherence  
Een jong meisje in V.A. D. uniform liep licht over het gazon.

Discourse marker: the conjunction or linking word expresses a strange relationship

Discourse marker | Coherence      Discourse marker | Coherence  
Inderdaad, ik had hem nog nooit zo goed gekend.

Discourse marker | Coherence  
Hij was ruim vijftien jaar ouder dan ik, in de eerste plaats omdat hij er amper vijfenveertig jaar uitzag.

Co-reference: co-reference problem, mismatch between entities, e.g. feminine pronoun to refer to a male person

Logical problem | Coherence  
Ik was buiten het front geïnvaleerd; en na een aantal maanden in een nogal  
Logical problem | Coherence      Co-reference | Coherence  
deprimerend herstellend huis te hebben doorgebracht, kreeg hij een maand ziekteverlof.

Maar de Logical problem | Coherence materie was hem meteen Logical problem | Coherence aangegrepen , nam hem aan als Co-reference | Coherence secretaresse ... je weet hoe ze altijd honderd Logical problem | Coherence samenlevingen bestuurt? '

Inconsistency: a term or a notation is used inconsistently throughout the text

Het dorp Styles St. Mary lag ongeveer drie kilometer van het kleine station  
Ze werkt in het Rode Kruis-ziekenhuis in Tadminster, zeven Inconsistency | Coherence mijl verderop.

" Style & Register | Disfluent sentence/construction Waar is thee vandaag ... binnen of buiten?"  
"  
Inconsistency | Coherence Uit .

Verb tense: the tense of the verb is wrong/illogical in the context of the rest of the sentence/text

wilde, ongetemde geest in een prachtig geciviliseerd lichaam ... al deze dingen  
Verb tense | Coherence worden in mijn geheugen gebrand.

## Lexicon

One lexical element does not entirely fit in the Dutch sentence.

Lexical choice: the meaning of the word is clear in the context, but it is not idiomatic in Dutch

Lawrence, de Lexicon | Lexical choice jongere , Lexicon | Lexical choice was (flu) Lexicon | Lexical choice een delicate jeugd geweest .

Lexicon | Lexical choice Ze wierp zich naast John op de grond en terwijl ik haar een bord boterhammen overhandigde, glimlachte ze naar me.

Lexicon | Lexical choice een grondig fascinerende vrouw.

Wrong preposition: this expression requires a different preposition

Lexicon | Wrong preposition Ik zag meteen door zijn gezicht dat er iets verontrustends had plaatsgevonden

De intense interesse die Lexicon | Wrong preposition in het publiek werd gewekt door wat destijds bekend stond als 'The Styles Case' is nu enigszins verdwenen.

Non-existing/foreign word: the word does not exist in Dutch and makes no sense

Mevr. Cavendish was echter een dame die graag haar eigen plannen maakte, en verwachtte

dat andere mensen bij hen Logical problem | Coherence zouden vallen, en in dit geval had ze zeker de

Non-existing word | Coherence zweefhand, namelijk: de Non-existing word | Coherence buidelrnen.

"En ze is doodsbang voor ' Style & Register | Untranslated Sisters '."

## Grammar & syntax

Anything that does not follow the grammatical or structural rules of the Dutch language

Agreement: The sentence is grammatically incorrect because there is a mismatch between and article and noun, subject-verb, etc.

Grammar & Syntax | Agreement Logical problem | Coherence Grammar & Syntax | Agreement  
Dit vertrouwen zullen de sensationele geruchten stoppen die nog steeds bestaan.

Verb form: a grammatically incorrect verb form

Word order: a grammatically incorrect word order

Grammar & Syntax | Word order Grammar & Syntax | Word order  
"Geen geheime hobby?" zij vroeg.

Extra words: The sentence is grammatically incorrect because there is a superfluous constituent or word. This can be a constituent that already appeared earlier in the sentence or a word that has been written twice in a row

Het dorp Styles St. Mary lag ongeveer drie kilometer van het kleine station en  
Grammar & Syntax | Extra word  
het Styles Court lag anderhalve kilometer aan de andere kant ervan.

Missing words: The sentence is grammatically incorrect because a necessary part of the structure is missing, this can be a preposition or an entire constituent such as an obligatory direct object.

Grammar & Syntax | Missing word

Daarna wendde zich tot zijn vrouw:

Grammar & Syntax | Missing word

Miss Howard schudde de hand met een stevige, bijna pijnlijke greep.

Grammar & Syntax | Missing word

Ik sprak mijn waardering uit, en John bracht me naar het huis en de brede trap op, die rechts en links voorwaarts slingerde naar verschillende vleugels van het gebouw.

Other: Other grammatical error that does not belong to any of the previous categories.

## Style & register

Disfluent sentence/construction: The sentence / constituent is not grammatically incorrect, but it is nonetheless very difficult to read, it could be translated in a much more idiomatic way.

Style & Register | Disfluent sentence/construction

Ik had een paar jaar heel weinig van hem gezien .

Style & Register | Disfluent sentence/construction

Mevr. Cavendish, die met zijn vader was getrouwd toen hij weduwnaar met twee zonen was,

Style & Register | Disfluent sentence/construction

was een knappe vrouw van middelbare leeftijd geweest, zoals ik haar herinnerde .

Repetition: The same or a very similar word/expression is used more than once

Style & Register | Repetition

en een knappe oude oude dame

Style & Register | Repetition

Cynthia Murdoch was een jong uitzierend jong wezen, vol leven en kracht.

Register: The words have the same meaning, but the chosen word/expression is too formal/informal/... for the text or belongs to a regional variety of the language that is not entirely suitable for the target audience.'

Orthography: Errors according to the methodology of writing Dutch language.

Spelling: wrong spelling of a word

"Pesten ze je dan?" Spelling Vroeg ik glimlachend.

Een dame in een stevige Spelling tweedrok,

Capitalization: Incorrect capitalization

Punctuation: Incorrect/additional/missing punctuation mark(s).

Other: Other orthography error that does not belong to any of the categories above.

Multiple errors: A combination of fluency errors which make it difficult to annotate the errors separately.

NL: zodat ze kunnen het opzetten van kleine bedrijven  
-> rephrase as "zodat ze kleine bedrijven kunnen opzetten"

NL: Sneeuwval vertraagt het drukke verkeer op de snelweg tussen Bratislava en Brno, maar we nog steeds lukt ...  
-> rephrase as "het lukt ons nog steeds"

NL: Daarom werken we aan een betere dienstverlening door het versterken van onze supply chain.  
-> rephrase as "onze productieketen te versterken"

Other: Other fluency error that does not belong to any of the categories above.

## Step 2: Accuracy errors

### Accuracy

Mistranslation	Do Not Translate	Untranslated	Addition	Omission	Mechanical	Terminology
Multiword Unit (MWU)					Capitalization	
Word Sense					Punctuation	
Semantically Unrelated					Other	
Part-of-Speech (POS)						
Partially translated						
Other						

The goal of the second step is to annotate the machine translated text for accuracy. Accuracy can be described as making sure the target text contains the same information as the source text. This means that all misinterpretations, contradictions, meaning shifts, additions or deletions are potential errors.

1. Before you start annotating for accuracy, make sure the Accuracy layer is selected in Webanno.
2. Accuracy annotations require two sets of annotations (one in source sentence and one in target sentence) and a "link" from source to target annotation.
3. First annotate the source text error, then annotate the corresponding translation in target text. Link annotation in the source text to the annotation in the target text with an arrow.

Accuracy has 6 main categories to choose from: Mistranslation, Do Not Translate, Untranslated, Addition, Omission and Mechanical.

Mistranslation has subcategories, namely Multiword Unit, Word sense, Semantically unrelated, Part of Speech, Partially translated. Mechanical errors also have subcategories, namely Punctuation, Capitalization and Other. The difference between Mechanical errors (in Accuracy) and Orthography errors (in Fluency) is that while the fluency errors can be detected (and related) to the target language alone, accuracy errors refer to transfer errors from source to target.

## Mistranslation

Source content has been translated (when it should be translated) but the translation is incorrect. Choose the suitable subcategory.

Multiword Unit (MWE): The translation is incorrect (and often too literal) because the English sentence contained a multi-word expression such as an idiom, a proverb, a collocation, a compound or a phrasal verb. Idiomatic expressions and proverbs are indicated with a paragraph sign (§) in Van Dale.

The screenshot displays a Webanno interface with two sentences for annotation. The top sentence is in English: "Mrs. Cavendish, however, was a lady who liked to make her own plans, and expected other people to fall in with them, and in this case she certainly had the whip hand, namely: the purse strings .". The bottom sentence is in Dutch: "Mevr. Cavendish was echter een dame die graag haar eigen plannen maakte, en verwachtte dat andere mensen bij hen zouden vallen , en in dit geval had ze zeker de zweefhand , namelijk: de buidelrnen .".

Annotations are shown as boxes with labels like "Mistranslation | MWE" and "Non-existing word | Coherence". Red arrows labeled "(acc)" connect the English annotations to the Dutch ones. For example, "fall in with" is linked to "bij hen zouden vallen", "had the whip hand" is linked to "de zweefhand", and "purse strings" is linked to "buidelrnen".

*to fall in with* = zich aansluiten bij  
*to have the whip hand* = macht hebben over  
*purse strings* = touwtjes van de geldbeurs

But you could have Mistranslation | MWE knocked (acc) Mistranslation | MWE us all down with a feather when, three months ago, she suddenly announced that she and Alfred were engaged!

(acc) Mistranslation | MWE  
Logical problem | Coherence

Maar je had ons allemaal met een veer kunnen slaan toen ze drie maanden geleden plotseling aankondigde dat zij en Alfred verloofd waren!

*knock down with a feather* = zeer verbaasd zijn

Word sense: The target content refers to a different (and a wrong) sense of the source content.

Mistranslation | Word Sense

A great sport ...old Evie!

(acc) Mistranslation | Word Sense  
Logical problem | Coherence

Een geweldige sport ... oude Evie!

*sport* = meid (Brits-Engels; informeel)

Mistranslation | Word Sense

Come and be refreshed ."

(acc) Mistranslation | Word Sense  
Logical problem | Coherence

Kom en word vernieuwd ."

*refresh* = zich opfrissen

Mistranslation | Word Sense

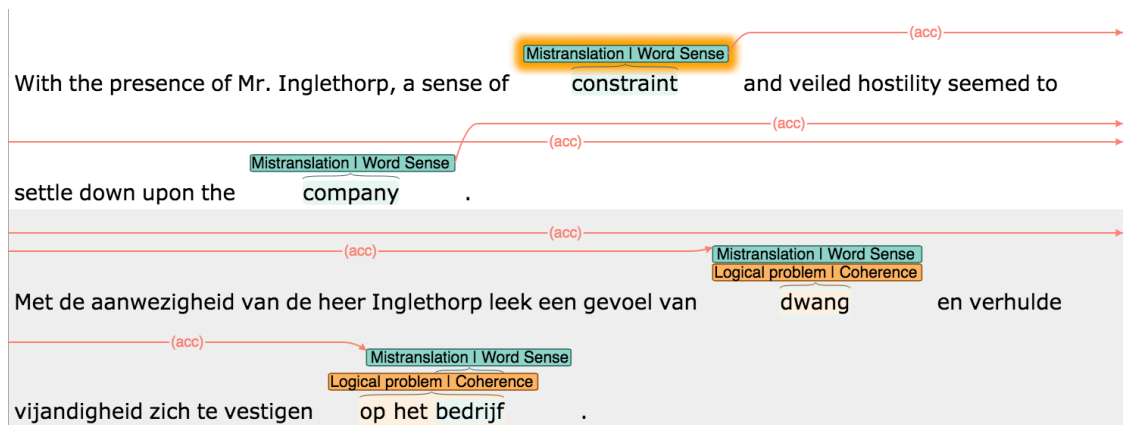
A man followed her, a suggestion of deference in his manner.

(acc) Mistranslation | Word Sense  
Logical problem | Coherence

Een man volgde haar, een suggestie van eerbied op zijn manier.

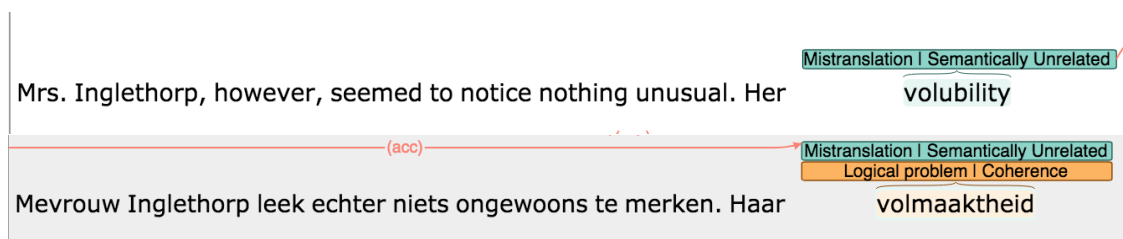
*suggestion* = spoor, zweem





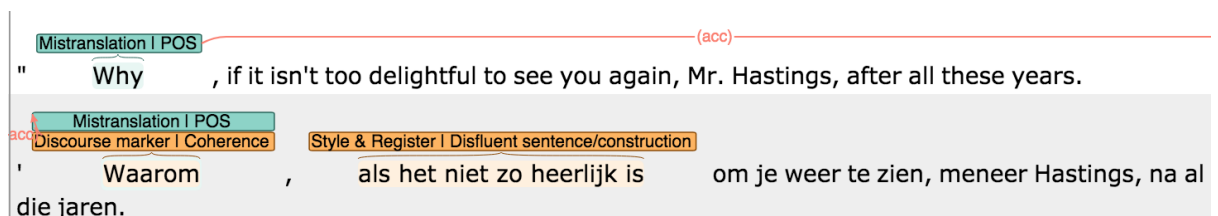
*constraint = verkrampd gedrag*  
*company = gezelschap*

Semantically unrelated: The meaning of one or more translated words is not related in any way to the meaning of the source word(s) and does not make any sense in the context.



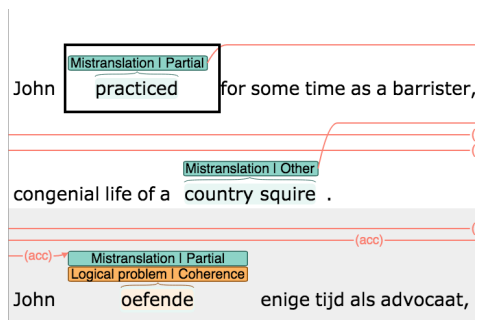
*volubility = welbespraaktheid*

Part-of-Speech: The translation represents an incorrect lexical category (Part-of-Speech) of the corresponding source text.



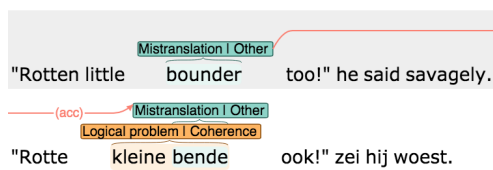
*why = tussenwerpsel hier: wel*

Partially translated: The translation is incorrect, e.g. due to the partial translation of a Dutch separable verb.

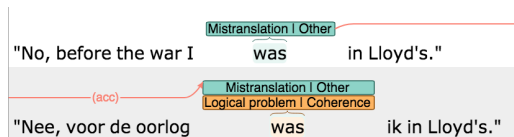


*practice = een beroep uitoefenen*

Other: None of the categories above



*bounder = onbeschoft persoon*



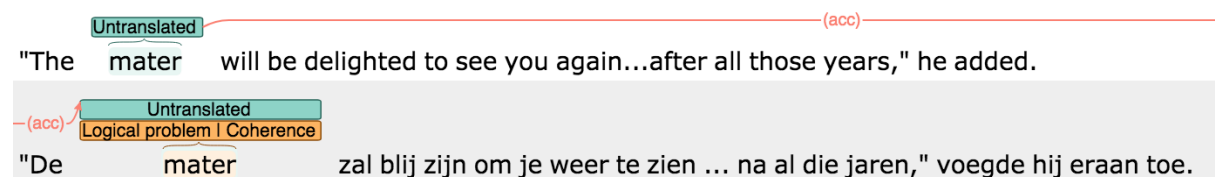
*was = werkte ... bij*

## Do Not Translate

The source content is unnecessarily translated into target language when it should have been left untranslated.

## Untranslated

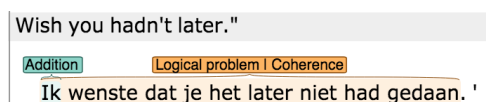
The source text is not translated (but was copied to target) when it should have been translated into Dutch.





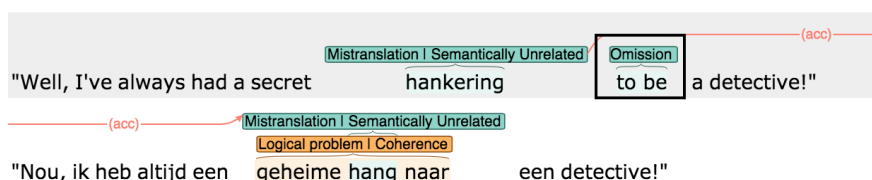
## Addition

Target text that is not present in the source sentence.



## Omission

Source content that cannot be found in target sentence.



## Mechanical:

Error related to incorrect transfer of capitalization, punctuation (or other orthographical features) from the source text to the translation.