

How to write for good machine translation quality

A controlled language is an interesting solution for authors who write texts for translation.

A controlled language has stricter rules than the general language. Controlled languages enable authors to write texts that are easily comprehensible. And it is an undeniable fact: Texts that are easily comprehensible are also easily translatable. Human translators and machine translation systems produce much better results if authors write the source texts in a controlled language.

A controlled language with ten rules.

Each natural language has its own grammar. Therefore, the rules for controlled languages differ from language to language. In other words: There can be no rules that produce optimal results for all languages. But the following rules will reduce ambiguities in most texts in many languages. And texts without ambiguities are ideal for machine translation.

These controlled language rules are examples of the rules in the CLOUT™ rule set. CLOUT is an acronym for *Controlled Language Optimized for Uniform Translation*. The CLOUT™ rule set was developed by Uwe Muegge.

Controlled Language Rule 1:

Write sentences that are shorter than 25 words.

Write:

The author performs the following tasks:

- ◆ Collect the necessary information.
- ◆ Analyze and evaluate the information.
- ◆ Write a structured draft.

Do not write:

Authors will approach any writing project by collecting the necessary information first, and after carefully analyzing and evaluating it, they will create a structured draft.

Controlled Language Rule 2:

Write sentences that express only one idea.

Write:

Authors who optimize their texts for easy comprehension facilitate the translation process. These texts enable human translators and machine translation systems to produce better translation results.

Do not write:

By optimizing their texts for easy comprehension, authors facilitate the translation process, and doing so enables human translators and machine translation systems to create better translation results.

Controlled Language Rule 3:

Write the same sentence if you want to express the same content.

Write:

Printer Installation.

- 1) Remove the printer from the shipping carton.
- 2) Remove the printer from the plastic bag.

...

Scanner Installation.

- 1) Remove the scanner from the shipping carton.
- 2) Remove the scanner from the plastic bag.

Do not write:

Instructions for installing the printer.

After unpacking the printer from the shipping carton, take the printer out of the plastic bag.

...

How to install your scanner.

- 1) Lift the scanner from the shipping box.
- 2) Discard all packaging material.

Controlled language rule 4:

Write sentences that are grammatically complete.

Write:

Do you wish to continue the installation of the software?

Do not write:

Continue installing software?

Controlled language rule 5:

Write sentences that have a simple grammatical structure.

Write:

Show that you can organize your thoughts by using a simple sentence structure in your texts.

Do not write:

You, in your texts, to show that you can organize your thoughts, should use a simple sentence structure.

Controlled language rule 6:

Write sentences in the active form.

Write:

The program manager will send a summary of questions to the responsible individuals.

Do not write:

A summary of questions will be sent to the responsible individuals.

Controlled language rule 7:

Write sentences that repeat the noun instead of using a pronoun.

Write:

You must check the spelling of your text before you publish your text.

Do not write:

You must check the spelling of your text before publishing it.

Controlled language rule 8:

Write sentences that use articles to identify nouns.

Write:

Test the installation.

Do not write:

Test installation.

Controlled language rule 9:

Write sentences that use words from a general dictionary.

Write:

Avoid ambiguity.

Do not write:

Eschew obfuscation.

Controlled language rule 10:

Write sentences that use only words with correct spelling.

Write:

Texts that contain spelling mistakes complicate the translation process.

Do not write:

Texts that contain spelling mistakes complicate the translation process.

A controlled language is particularly useful for certain text types.

The following text types are particularly suitable for the use of a controlled language:

- ◆ Software strings.
- ◆ Help systems.

- ◆ Technical specifications.
- ◆ Technical reports.
- ◆ Technical documentation.

A few examples of organizations that use a controlled language.

This list provides the name of a few organizations that use a controlled language. This list also provides the name of the controlled language that that organization uses.

- ◆ Avaya: Controlled English (ACE)
- ◆ Boeing: Simplified Technical English (STE), ASD-STE100.
- ◆ Caterpillar: Caterpillar Technical English (CTE), Caterpillar Fundamental English (CFE).
- ◆ Dassault Aerospace: Français Rationalisé.
- ◆ European Aeronautic Defence and Space Company (EADS): Simplified Technical English (STE), ASD-STE100.
- ◆ Ericsson: Ericsson English.
- ◆ General Motors (GM): Controlled Automotive Service Language (CASL).
- ◆ IBM: Easy English.
- ◆ Kodak: International Service Language.
- ◆ Nortel: Nortel Standard English (NSE).
- ◆ Océ: Controlled English.
- ◆ Rolls-Royce: Simplified Technical English (STE), ASD-STE100.
- ◆ Saab Systems: Simplified Technical English (STE), ASD-STE100.
- ◆ Scania: Scania Swedish.
- ◆ Sun Microsystems: Sun Controlled English.
- ◆ Xerox: Xerox Multilingual Customized English