The Janus Head Article - On Quality in the Documentation Process - By: HENRIK ANDERSEN, ELLEN CHRISTOFFERSEN, & MARGRETHE H. MØLLER
Posted by Michael Lambarena on February 22 2009 13:45:38

The Janus Head Article

- On Quality in the Documentation Process

The god Janus in Greek mythology was a two-faced god; each face had its own view of the world. Our idea behind the Janus Head article is to give you two different and maybe even contradicting views on a certain topic. In this issue the topic is quality in the documentation process. In the first half of this issue's Janus Head Article translators from the international company Grundfos give us their view of quality and how quality is managed in the documentation process at Grundfos. In the second half of the Janus Head Article scholars from the University of Southern Denmark describe and discuss quality in the documentation process at Grundfos from a researcher's point of view.

Controlled Language - An Issue for Grundfos?

The following articles discuss the issue of "Controlled Language for Danish Enterprises" viewed from the perspectives of two project partners from both practice and research. One is the Danish multinational pump manufacturer Grundfos and the other is a research group from the University of Southern Denmark.

Grundfos is a pioneer in the area of language policy and is continuously striving to improve both its documentation process and the quality of its documentation. In the article, Grundfos describes and discusses its efforts to develop and improve its overall documentation process and the tools, people and organisation involved.

The research group based at the University of Southern Denmark, has a long history of both researching and teaching in language technology. In the article, they discuss their analyses and first recommendations for Grundfos within the research project.

Finally, Grundfos comments on the findings in a conclusive

statement.

- the Grundfos Perspective

The main purpose of this article is to address the documentation process at Grundfos related to technical documentation. More specifically, the technical documentation developed and produced by the department called Technical Marketing. Technical Marketing produces material such as product descriptions, data booklets (catalogues), installation and operating instructions, and service instructions.

Over time, we have focused our efforts on improving both the documentation itself and the documentation process. This has mainly involved one overall process on the strategic level The Documentation Process - and three tactical/operative aspects: Tools, People and Organisation.

Documentation Process

The documentation process is a support process, or a generic model, assisting Grundfos staff in developing and producing documentation. The process was developed as a joint venture among several departments (as illustrated in fig.1: The Grundfos Documentation Process).
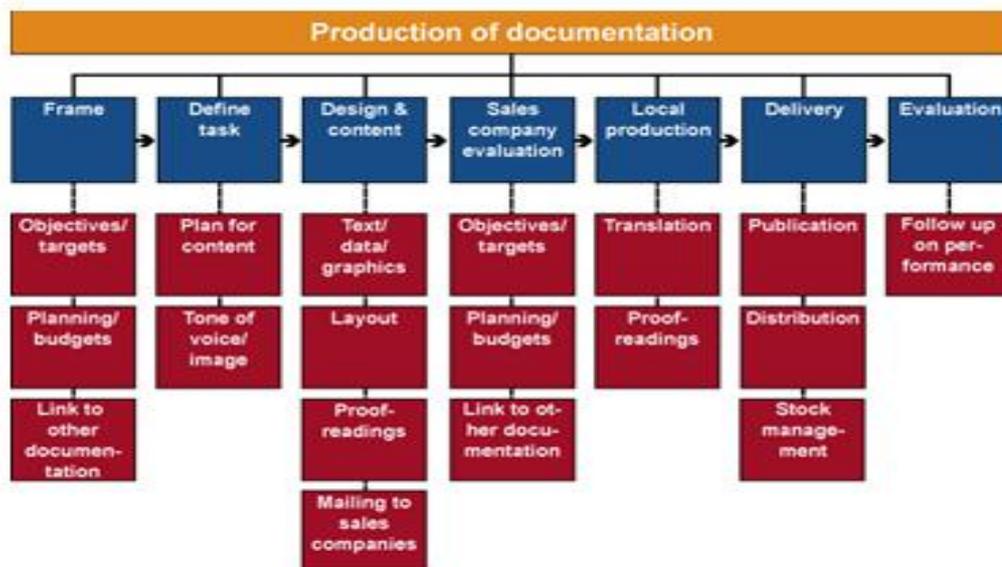
Grundfos' Documentation Process

As indicated, the model is made up of several steps that one must go through to produce good documentation, i.e. from definition of

task, production of document, translation to evaluation.

It may be appropriate to mention here that it is Technical Marketing's policy to write the basic text version in Danish. The background of this policy is quite simple: All staff in Technical Marketing are Danes and therefore master the Danish language far better than any foreign languages. The translation into English - which is the basis for translating into other languages - is then done by skilled, in-house translators who are experts in translating technical texts.

The rest of the article will focus on the tactical/operative aspects:

• Tools - covers a number of computer programs with the aim of improving or speeding up the work-flow.
• People - means enhancing the competence level of staff through training and education.
• Organisation - relates to various organisational adjustments, both in terms of optimising the work processes and the physical office layout.

Tools

MultiTerm

MultiTerm is the name of Grundfos' electronic dictionary.

MultiTerm

• was compiled using the corpus-concept, meaning that the entry words were extracted from two large corpora of Danish texts and their English translations.
• was based on the lexicographic concept with English as the head-term, i.e. English synonyms are generally not accepted in the same entry.
• (English was chosen as the head-term language in order to prepare MultiTerm for the possible inclusion of other languages at a later stage)
• included Danish equivalents, grammar, context, etc. (Over the years, German equivalents have been added to many of the entries).

The establishing of MultiTerm had several purposes:

• to retain terminological knowledge in the department that would otherwise be lost when staff leave.
• to share and enhance the common terminological knowledge.
• to ensure consistency in the use of terms - for the benefit of the

readers and to enhance the general level of quality.

• to save time otherwise spent on searching for terms.

MultiTerm has been used and further developed since 1996 and today has almost 6,000 entries.

Translation Memory

In 2000, it was decided to introduce Trados' Translator's Workbench into Technical Marketing.

The aim of introducing the Workbench was more or less the same as for introducing MultiTerm, with the added feature that now we could also save and share linguistic and syntactical knowledge. Since the introduction, development has been fast and we have seen a remarkable increase in translation productivity and quality (consistency) as a result.

WebTranslate

The obvious benefits of a translation memory led to strategic considerations and - in the end - a decision by Technical Marketing to develop an in-house translation management program, called WebTranslate. The advantages of developing our own program, instead of expanding the use of the above-mentioned translation memory, were numerous.

Developing our own program enabled us to:

• meet exactly our own defined requirements, such as capability of handling unicode character sets.
• co-operate very closely with the developer, enabling precise adaptation to our needs and - after development - ensure running adjustments and improvements to keep the program up-to-date with our needs.
• utilise the benefits of running the program on the Web (which was not very common when we started out).
• train both internal (in Grundfos' sister companies) and external users in a targeted manner.
• incorporate a project management module adapted precisely to our needs.
• take on competition with the professional translation memory developers (all the know-how collected in the process of developing the program has later been transferred to an independent enterprise).

Apart from this, we can also capitalise on both the common and well-known benefits of working with a translation memory, as we can:

- save time and resources - meaning a reduction in the time to market, thus, resulting in pay-back on the investment.
- save and share linguistic and terminological knowledge, thereby enhancing quality and consistency of language and terminology (thus, strengthening Grundfos' competitive edge and image).
- increase productivity and job satisfaction.

WebTranslate has now been implemented full-scale. This means that it is running in 26 Grundfos companies all over Europe as well as with some external partners. All receive the same basic text in English uploaded on the Web, translate the text into their local language, then proof-read and approve the final text. Translation project management, as well as program adjustments and improvements lie in Technical Marketing. The preliminary evaluation after the first 12 months is very positive as problems grow smaller and smaller and productivity increases steadily.

WebTranslate Term Base

The next phase will be to convert our existing MultiTerm database to WebTranslate. The idea is to provide WebTranslate with a term-recognition facility, meaning that apart from hits for sentences or translation units, the translator will also be given suggestions for the translation of individual terms. In the long run, the WebTranslate term-base may be opened to a wider audience, thus making internal terminology and specialised term knowledge accessible to Grundfos staff all over the world.

The aim is of course, to streamline the terminology used in Grundfos texts through a higher degree of consistency.

Controlled Language

We are currently cooperating with researchers from the University of Southern Denmark on the project, "Controlled Language for Danish Enterprises", in order to find out whether a controlled-language checker could be introduced as an additional tool, in order to enhance quality in the writing phase. This cooperation and its preliminary results are discussed below in the article " -the research perspective".

People

Language Policy and Language Guides

In its Group Policy for Communication, Grundfos states that "High quality is the only acceptable level of language and communication". This statement was the direct inspiration of a project aimed at creating a language guide for the Danish language. The aim of the five-member project group was to formulate a set of general guidelines and attitudes related to the writing of Danish in the Danish Grundfos companies. The Danish language guide has three points of focus:

1) Useful considerations related to **the writing process**. Among these are: Who am I? Who is the reader? What is my message?

2) **The text itself**. The advice is to write short sentences, always use the same word for the same thing, adapt the text to the reader, create flow in the text, etc.

3) **Proof reading**. Never go lightly through proof reading. Spend the necessary time on doing a good job.

After the Danish language guide was introduced, the project group was asked to write a counterpart for Grundfos' corporate language, English. For this purpose, the project group was extended by members from four large Grundfos sister companies. The result of the work was the English language guide we called, "In other words...".

"In other words ..." focuses on the same issues as the Danish language guide (the writing process, the text itself and proof reading). However, the English language guide has been improved and extended considerably on several points, such as, how to write for electronic media and how to allow for cultural differences.

Even though we have not recorded any directly measurable improvements of the general quality of language in communication among Grundfos companies as a result of our language guides, there is no doubt that they have helped put the spotlight on the quality of language as an important part of our company's total image.

Focus on Communication in Technical Marketing

Communication of product knowledge in a plain and unambiguous language is the key competence in Technical Marketing.

Consequently, we have held a number of seminars and workshops for all writing members of Technical Marketing's staff. These workshops have all aimed at improving our common skills in writing a clear, plain and straightforward language.

Our ambition is that the reader understands the message the first time he reads a text. Focus has been on trying to put ourselves in the reader's position, asking all the questions that we can imagine a reader would ask. Next, we have concentrated on the writing process which we have divided into three phases: Before, During and After.

Before (writing): Picture the audience, make clear what the message is going to be, collect information and data, organise the information in a rough structure.

During (writing): Write the core-message first and leave all the explanatory and causal text until afterwards, write only one piece of information per sentence, use short words and short sentences, be precise (do not leave the reader with unanswered questions), make sure the sentences are linked together, make sure headlines and figures match the body text.

After (writing): Be very thorough when proof reading, check and double-check information, structure and language, read the text aloud, let a colleague proof read your text.

For the above-mentioned purpose, a very creative project group in Technical Marketing created our own little writing guide called "The good HABIT", in the form of an A4-page folded and glued to have three sides. On each side you can see the central advice of what to do before, during and after writing the text. You can stand the three-sided A4 on your desk so that you always have the relevant side facing you.

Technical Marketing's writing guide, "The good HABIT"

All this may seem simple and self-evident, however, when we started the process we were a bit embarrassed to see how often we "forgot" about these simple rules. Only when we put the spotlight on them and kept our focus concentrated on them, did we achieve a measurable improvement. Since then we have regularly sent texts for external evaluation, and are proud to see that we have received better and better results every time. On a scale of 100 percent, we pride ourselves on having reached 90 percent, thus, signalling that there is room for improvement. However, as the remaining 10 percent all lie in the minute details, our consideration now is how we reach 100 percent without spending disproportionate time and resources.

Future Focus Areas in Technical Marketing

Improving the quality of documentation is an on-going process in Technical Marketing. Every year has its own game plan of projects and initiatives. This year, the game plan includes: projects to define and measure quality, improve the proof reading process and optimise document contents, just to mention a few projects that are directly related to communication. We see the research project in controlled language as an interesting, supplementary project.

Organisation

Project Teams

Up until 2004, the internal Technical Marketing department was divided into groups based on people's educational background. Technicians and engineers formed one group, responsible for all products and document types, and the linguists formed a translators' group, responsible for finishing the Danish draft and translating all texts.

In 2004, these two groups were reorganised to form four project teams, each consisting of two or three technicians/engineers and one translator. Each project team is responsible for a defined product portfolio and for all the documents to be produced on those products.

The purpose of the reorganisation was:

1) to involve the translators at a much earlier stage in the writing process, enabling them to provide their high-level linguistic competence already in the original writing phase.

2) to give the teams a higher degree of ownership of the documents concerning their product portfolio.

The project teams are responsible for content, language and layout of documents, and in this work, they are assisted by a number of support functions:

Illustrators: Draughtswomen managing the illustrative work, such as product drawings, sketches, diagrams and photos.

Curves and data: Staff taking care of drawing the performance curves and of managing all the many facts-related data, such as dimensions, weights and electrical data.

DTP: Desk-top publishing staff responsible for the final document layout and of coordinating translations into all languages other than English. DTP also manage the ordering procedures related to our printing house.

The reorganisation and the increased focus on communication were parallel developments. And this has contributed greatly to the fact that - in our opinion - we are getting closer and closer to an ideal text production process. Through very close cooperation and many feedback loops, we utilise the skills and competence of the various

groups of employees in the best possible way. We are convinced that this - combined with our open-plan office environment where project and support teams work in the same room - has contributed to the quality of our documents.

Conclusion

As it appears, developing and improving document and process quality is an on-going process where focus must be on creating the optimum synergy among the three central aspects: Tools, People and Organisation. It is highly evident that the process is dynamic and that technological advances and new know-how, as well as user expectations will constantly force us to move forward. The challenge is to find the right balance so that we do not lose sight of the main objective: To write the best technical documentation in the world.

- the Research Perspective

What is Controlled Language?

-A controlled language (CL) is a set of writing rules specifying constraints on lexicon, grammar, and style. It could be described as a very specific style-guide. The CL-lexicon specifies approved terms, non-approved synonyms, definitions and examples. The CL-rules specify approved and non-approved sentence structures. See examples in Appendix A.

-CLs are normally implemented in specialised term, grammar and style-checkers tuned to the writing rules of the company, so-called CL-checkers.

-CLs are used primarily for technical texts, such as user manuals, and although most CLs share a common core, they need to be tailored for each organisation or enterprise.

-The objective is to improve text quality. Ambiguity and complexity are reduced with a view to making texts easier to read, understand and translate.

-As an additional benefit, controlled texts will increase the efficiency of language technology. In translation memory systems, the number of hits (the leverage) is increased, and in machine translation systems, the quality of raw translations is improved.

-Most CLs have been designed for English. Most CL users are major companies such as XEROX and General Motors. The aviation industry has agreed on a common standard for Simplified Technical English. ASD-STE100[TM].

In the project "Controlled Language for Danish Enterprises", we are investigating methods for designing controlled languages (CLs) for Danish and testing these methods on Danish enterprise texts.

Another objective of the project is to develop Danish modules for a controlled language checking program, a so-called CL-checker.

In the project, we cooperate with several enterprises, one of which

is Grundfos.

This article is based on the first phase of our cooperation with Grundfos, where a quality analysis of Grundfos' technical documentation was carried out and a prototype of a controlled language for Grundfos was developed.

Grundfos has a reputation for focusing on documentation quality. This makes Grundfos an ideal project partner for us, as the company has a formulated language policy and a term-base which may, in theory, be used as a basis for grammatical and stylistic CL-rules and for a CL-lexicon. Furthermore, Grundfos' documentation is translated by way of translation memory systems. This means that Grundfos has a double interest in controlled language - not only is it important that the documentation is well written and easy to read, it must also be highly standardised so as to yield a high leverage in a translation memory.

The figure below illustrates where a Danish CL-checker would fit into the Grundfos Documentation Process: It could be used by the authors of the Danish source texts and/or by the proofreaders.
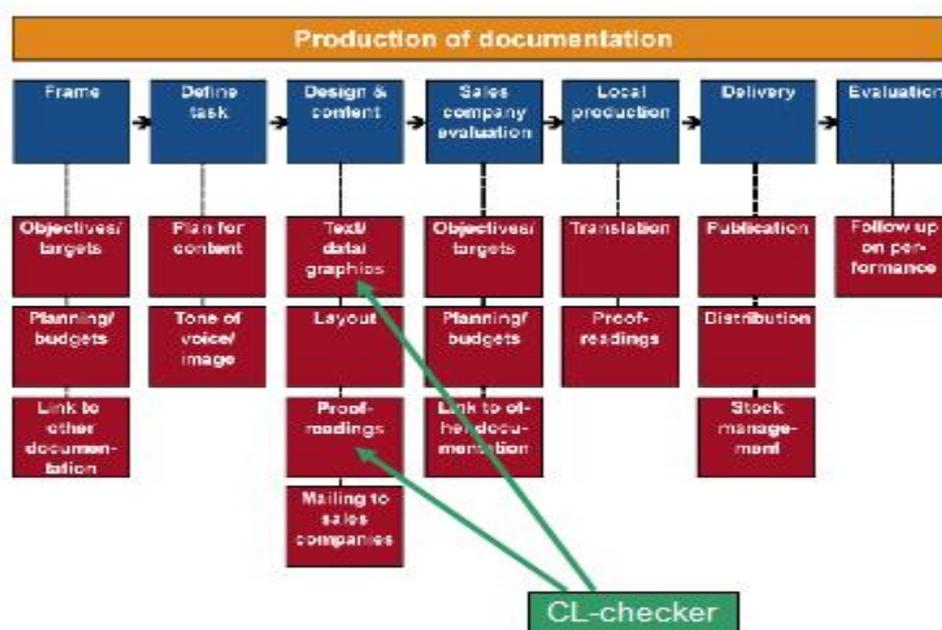
The place of a CL-checker in the Grundfos Documentation Process

Grundfos made a number of texts, as well as their MultiTerm term-base, available to us for our analyses and gave us feedback on our suggestions for CL-rules and CL-lexicon. The texts consist of technical manuals: a large corpus of publicised texts and a minor

corpus of draft texts.

The Publicised Texts

The publicised texts clearly reflect the fact that Grundfos has a language policy and that much effort is put into proofreading. In other words, the texts are of high quality, and there are only few problems relating to information structure, sentence structure and punctuation.

As we see it, the problems are mainly related to terminology. According to the literature on controlled language, and in terminology theory, the guiding principle is "one word - one meaning". Likewise, the Grundfos style-guide recommends using the same word for the same thing. However, living up to this ideal is not easy and consequently, the publicised Grundfos manuals have examples of:

a) several designations being used for one-and-the-same thing (synonyms and orthographic variants, e.g. Danish akseltap and akselende, both meaning shaft end) and

b) one designation being used about different things (homonyms, e.g. Danish alarm meaning either alarm signal or alarm unit).

Both types of ambiguity may confuse the reader. Likewise, ambiguities make the texts more difficult to translate by means of language technology tools. Synonyms and orthographic variants make the texts less standardised and thus, less suited for translation by means of a translation memory system. Homonyms make the texts more ambiguous and thus, less suited for translation by means of a machine translation system, as it may be difficult for the system to find out what the word or sentence in question means and, therefore, how it should be translated.

Analysing the Grundfos texts, we found that it would be possible to eliminate most synonyms and orthographic variants.

On the other hand, we found that, in practice, it would not be possible to eliminate the majority of the homonyms occurring in the texts. This is surprising, as the principle of "one word - one meaning" has hardly ever been questioned in the literature (as most controlled languages are company-specific and confidential, we do not know what is done in practice).

There are several reasons why homonymy is difficult to eliminate. Firstly, many words and terms have different meanings within

different subject areas. Technical documentation often includes more than one subject area, and in practice, it would not be possible to replace an established term with a new term in order to avoid homonymy. Secondly, homonymy is often caused by compression. When a term such as cooling system or alarm system has been introduced in a user manual, it will typically be referred to by means of the head of the compound term, i.e. system, in the rest of the paragraph or chapter, and only the context will clarify which type of system is meant. A CL-rule banning compression would result in stilted texts.

Nevertheless, a controlled language such as AECMA Simplified Technical English (now ASD-STE100TM) actually attempted to eliminate homonymy in order to prevent ambiguity. This was probably due to the intended audience - AECMA STE was designed for flight mechanics who were not native speakers of English. The Grundfos case is different, in that the intended audience of the Danish texts consists of engineers and pump fitters who have Danish as their native language. Therefore, homonymy and reduction are less problematic - also considering the fact that Grundfos uses a translation memory for translation, and not machine translation, as outlined above.

So, as far as the publicised Grundfos documentation is concerned, our conclusion was that there may be a need for a CL-checker which will flag non-approved Danish synonyms and orthographic variants, and suggest an approved term to use instead.

The question was whether information about approved terms and non-approved variants was already available in the Grundfos term-base. It turned out that this was not generally the case. The Danish part of the Grundfos term-base is primarily descriptive, i.e. it describes the Danish Grundfos vocabulary without systematically trying to prescribe a specific terminology, as a prescriptive term-base would. This is probably due to the fact that the primary purpose of the Grundfos term-base is to aid translation, not to standardise vocabulary. If the term-base were to be used for controlled language purposes, one task would be to decide which synonyms and orthographic variants should be chosen as approved terms. As the existing Grundfos term-base is not complete, in comparison to the Grundfos texts, another task would be to add missing, non-approved synonyms and orthographic variants and the corresponding approved terms to the term-base.

The Draft Texts

The problems mentioned above were found in the publicised texts. However, CL-checkers are not intended for texts that have already been proofread. They are meant as a support for the technical writer writing the texts, and/or the editor proofreading the texts.

Therefore, Grundfos made a number of draft texts available to us. In these texts, we found various spelling, grammatical and stylistical problems - including sentences which were completely garbled. In addition, there were sentence structures which were in conflict with CL-rules such as, "Avoid passive in instructions", "Avoid nominalisations", "Write only one piece of information per sentence in instructions, unless two actions are to be carried out simultaneously", "Avoid long adjectival phrases", etc.

In other words, there were two types of errors in the draft texts - general language problems and domain and company specific language problems. Instead of building a CL-checker capable of taking care of both types, one might consider a workflow where general problems were detected by a standard spell and grammar-check as offered, for example, by Microsoft Word, and only the domain and company specific language problems were detected by the specialised CL-checker. Therefore, we tested whether the Danish spell and grammar-check found in Microsoft Word 2003 was able to recognise the errors of the draft texts. It turned out that, in most cases, Word did not recognise the errors. The spell-check will only recognise spelling errors if the word in question is not identical to another correctly spelled word found in the spell-check lexicon. Danish compounds which were erroneously written with a space (e.g. *bus kommunikation instead of the correct spelling buskommunikation - a spelling error which occurs frequently in Danish texts because of the influence from English), are not recognised at all. The Danish grammar-check did not recognise any of the grammatical errors of the draft texts.

We concluded that a specialised CL-checker would be necessary in order to recognise the majority of the problem types found in the texts.

Which Tools?

A CL-checker works with linguistic knowledge. In order to complete its task, it must have a Danish analysis module consisting of both grammar and lexicon, and a Danish CL-module performing the language check, based on company-specific lexical, grammatical and stylistic rules.

One way of building a CL-checker for Danish would be through the construction of Danish modules for an existing CL-checker. This, however, is a very expensive project. None of the CL-checkers known to us cost less than 35,000 EUR. Add to this, the price for development of the above-mentioned Danish modules, which would probably amount to 1-2 person years. Some of the work to build the Danish modules may be part of a research project, however, the main development effort would have to be financed by external funds.

Denmark shares the conditions of other small countries, in that the market for language technology is limited, resulting in language technology suppliers having little interest in investing in Danish versions of their tools. Therefore, development of a Danish CL-checker will be dependent on funding by commercial partners as well as public funding. It is important to note, however, that once developed, the Danish analysis module, part of the CL-rules and some of the CL-lexicon, could be reused in other CL-applications.

One also has to ask oneself whether a CL-checker presents the full answer to the question put by Grundfos: How can we write good Danish source texts more efficiently?

As one important element of a controlled language is the checking of the terminology, Grundfos might consider to settle for a stand-alone terminology-checker, i.e. a program which is capable of recognising non-approved synonyms and orthographic variants in texts and suggesting the approved terms to use instead. However, the terminology-checkers we know of are components of, for example, CL-checkers, and not available as individual modules.

Another possibility might be to look at different types of authoring tools which are based on a principle of reusing sentences or text blocks.

One type of authoring tool works by comparing sentences to a translation memory as well as a term-base while the technical writer is writing, thus enabling him to adapt his source sentence immediately. The idea is to improve the consistency of source sentences and to increase the number of translations to be reused.

However, in the Grundfos texts, not only single sentences, but whole blocks of source text are reused. Therefore, another possibility might be an authoring tool which makes it possible to compose source texts from pre-defined building blocks.

These possibilities would have to be tested in the Grundfos documentation environment.

Concluding Remarks - by Grundfos

It is obvious that in our quest for excellence in text production, a kind of controlled-language checker would be attractive. However, a possible decision to push forward for such a tool will depend very much on the relationship between the investment and the estimated benefit. There is no doubt that time can be saved and quality enhanced, already in the writing phase, if a controlled-language checker assists the writer in choosing already approved terminology and linguistic structures. This advantage will be multiplied in the translation process, especially if the controlled-language database is coordinated with the translation-memory databases.

On the other hand, we are still considering the pros and cons. For some text types - typically, strictly instructive texts - a strict use of controlled language will be a clear advantage. However, there is no doubt to us that for descriptive texts, there is a limit to how much language can be "controlled" before you sacrifice your linguistic and stylistic freedom and a free and varied form of expression.

Authors

Henrik Andersen,

Senior Translator, Technical Marketing, Grundfos

Ellen Christoffersen,

Assoc. Prof.,

Institute for Business Informatics and Information Science,

University of Southern Denmark

Margrethe H. Møller,

Assoc. Prof.,

Institute for Business Informatics and Information Science,

University of Southern Denmark

References

Muldoon, Donna (1999): "CL 19: A Writer's View of Using a Controlled Language"

in: TC-Forum (Technical Communicator's Forum), www.tc-forum.org.

- Experience of a technical writer writing CL

Murphy, Dawn (2000): "Keeping Translation Technology under Control" in: Machine Translation Review Issue No. 11, December 2000 - pages 7-10, ISSN 1358-8346. Also published at http://www.bcs-mt.org.uk/mtreview/11/mtr-11-7.htm.

- Good presentation of the interplay between CL and TM and MT, respectively

Nyberg, Eric et al. (2003): Ch. 14 "Controlled language for authoring and translation" in: Harold Somers (ed.): Computers and translation: a translator's guide . John Benjamins B.V. 2003, ISBN 90 272 1640 1.

- The best and most detailed introduction to CL known to us

Appendix A

Examples of CL-lexicon entries (from Nyberg et al. 2003:246)

Approved word   prevent (v)

Definition   To make sure that something does not occur

Example   Attach the hoses to the fuselage to prevent their movement.

Unapproved word   preventive (adj)

Approved alternative prevent (v)

Unapproved example This is a corrosion preventive measure.

Approved rewrite   This prevents corrosion.

Approved word   right (adj)

Definition   On the east side when you look north

Example   Do a flow check of the pump in the right wing tank.

Unapproved word   right-hand (adj)

Approved alternative right (adj)

Unapproved example The fuel connector is in the right-hand wing.

Approved rewrite   The fuel connector is in the right wing.

Examples of Simplified English: prevent vs. preventive and right vs. right-hand

Examples of CL-rules of Simplified English (from Nyberg et al. 2003:247)

"Do not use sentences with more than 20 words"

"Do not use the passive voice"

"Do not make noun clusters of more than four nouns"

"Write only one instruction per sentence"

"Make your instructions as specific as possible"

"Use a bulleted layout for long lists"