

Chapter 1

Introduction to Global English

What Is Global English?	2
Why Global English?.....	2
Benefits of Global English for Professional Writers and Editors.....	3
The Cardinal Rule of Global English.....	4
Global English and Language Technologies	5
Machine-Translation Software.....	5
Translation Memory.....	7
Controlled-Authoring Software	8
Practical Considerations for Implementing Global English	10
Prioritize the Guidelines.....	10
Build a Relationship with Your Localization Staff.....	11
Eliminate Non-essential Information	11
Insert Explanations for Translators	13
Frequently Asked Questions about Global English	14
What is the relationship between Global English and controlled English?.....	14
Do the Global English guidelines make all sentences clear and easy to translate?	15

Does following these guidelines lead to an increase in word counts?	15
Typographical Conventions	16

What Is Global English?

In this book, Global English refers to written English that an author has optimized for a global audience by following guidelines that go beyond what is found in conventional style guides.

The Global English guidelines focus on the following goals:

- eliminating ambiguities that impede translation
- eliminating uncommon non-technical terms and unusual grammatical constructions that non-native speakers (even those who are quite fluent in English) are not likely to be familiar with
- making English sentence structure more explicit and therefore easier for non-native speakers (as well as native speakers) to analyze and comprehend
- eliminating unnecessary inconsistencies

Because Global English doesn't impose severe restrictions on the grammatical constructions or terminology that are permitted, it is suitable for all types of technical documentation.

Why Global English?

The Global English guidelines enable writers and editors to take the clarity and consistency of technical documents to a higher level, leading to faster, clearer, and more accurate translations.

Global English also makes technical documents that are not slated for translation more readable for non-native speakers who are reasonably proficient in English.¹ After all, many documents are never translated, and in today's world it is unusual for the audience of any technical document to consist solely of native speakers of English. Whether your audience consists primarily of scientists, engineers, software developers, machine operators, or unskilled workers, it probably includes a sizable number of non-native speakers.

Finally, Global English makes documents clearer and more readable for native speakers, too. Because native speakers of English still constitute the majority of the audience for many technical documents, that benefit should not be overlooked.

Depending on the type and subject matter of your documentation, Global English can provide the following additional benefits:

- Injuries, losses, and costly legal liabilities that can be caused by unclear documentation and by incorrect translations are avoided.
- Clearer, more-consistent documentation reduces calls to technical support.
- Consistent terminology facilitates the task of indexing and makes indexes more reliable.
- In online documentation or Help, users are better able to find the information they need because you have eliminated unnecessary synonyms and variant spellings.
- Translation quality is less of a concern because you have eliminated ambiguities and unnecessary complexities that can lead to mistranslations.

Benefits of Global English for Professional Writers and Editors

Anyone who produces technical documentation for a global audience should follow the Global English guidelines. However, many professional writers and editors have recognized two benefits of Global English that are less relevant to authors whose main responsibilities are in other areas.

¹ If your audience has limited proficiency in English, then consider using a form of controlled English in addition to following the Global English guidelines. See “What is the relationship between Global English and controlled English?” on page 14.

4 *The Global English Style Guide: Writing Clear, Translatable Documentation for a Global Market*

First, the ability to make documents more suitable for a global audience is a specialized and marketable skill. In a posting to a Society for Technical Communication mailing list, technical writer Richard Graefe made the following observation:

With the increase in localization² of documentation and of user interfaces, being able to sell yourself as a person with “pre-localization” editorial skills is a plus. To be able to do that type of editing well, you need . . . to be able to recognize English structures and expressions that will not translate well, that may be ambiguous to a translator, or that may require a translator to do rewriting in addition to translating.

The skills that Graefe described are exactly what this book helps you develop. In addition to giving job seekers a competitive edge, those skills could conceivably make professional writers and editors less vulnerable to layoffs and outsourcing.

Second, editors often find that the Global English guidelines articulate issues that they could not have explained themselves. The same is true for writers who are working with technical information that was provided by subject-matter experts. By referring authors to the explanations in this book, you can often persuade them to make the necessary changes with less resistance or discussion.

The Cardinal Rule of Global English

As noted above, native speakers of English probably constitute a significant portion of the audience for much of your technical documentation. Therefore, be sure to follow the cardinal rule of Global English even while you are taking into account the needs of non-native speakers and translators:

The Cardinal Rule of Global English

Don't make any change that will sound unnatural to native speakers of English.

² *localization*: the process of adapting products or services for a particular geographic region or market. Translation is a large part of the localization process.

At the same time, consider the following corollary to the cardinal rule:

Corollary

There is almost always a natural-sounding alternative if you are creative enough (and if you have enough time) to find it!

In other words, if following one of the Global English guidelines would cause a sentence to sound stilted or unnatural, then either find a different way to improve the sentence, or leave the sentence alone.



Consulting Colleagues

If you are a non-native speaker of English, your instincts about what sounds natural in English and what doesn't might not always be reliable. If you are not sure whether you are following the cardinal rule successfully, consult a native speaker whose judgment you trust.

Native speakers also benefit from consulting other native speakers on occasion. A colleague might quickly find one of those "natural-sounding alternatives" that eluded you.

Global English and Language Technologies

Often, people who are interested in Global English are also interested in technologies that make global communication more efficient. Three language technologies that are mentioned frequently throughout this book are machine-translation software, translation memory, and controlled-authoring software. The following sections provide an overview of these technologies and of how they relate to Global English.

Machine-Translation Software

Machine-translation (MT) software is software that translates sentences from one language (such as English) into one or more other languages (such as French or Japanese).

This book doesn't include guidelines that are useful *only* for improving the output of machine-translation software. However, many of the Global English guidelines that make documents more suitable for translation by human translators also make documents more suitable for machine translation. A sentence that is unclear, ambiguous, or otherwise problematic for human translators is often translated incorrectly by machine-translation software.

On the other hand, it is important to note that relatively little research has been done on the effect of specific style guidelines and terminology guidelines on machine-translation output. A guideline that improves the translated output for one MT system in one language might have a negligible effect (or, rarely, even a detrimental effect) for a different MT system or language.

Before implementing machine-translation software, ask the software vendor to help you identify and prioritize the Global English guidelines that would make your documentation most suitable for that particular software. Also consult sources such as Bernth and Gdaniec (2000), Roturier (2006), and O'Brien and Roturier (2007) to gain a better understanding of how to evaluate the effect of specific guidelines on MT output.

Case Study

A small pilot project conducted at SAS Institute indicated that the Global English guidelines have a significant effect on the quality of machine-translation output. This study did not examine the effect of specific guidelines. Instead, the project coordinators took a subjective look at the effect of following all of the Global English guidelines that were in the *SAS Style Guide for User Documentation* as of December 2004.

In this project, SYSTRAN translation software was used to translate the documentation for one software product from English to French. The process consisted of these steps:

1. Technical terms were pre-translated and added to the SYSTRAN dictionary.
2. A small part of the document was translated without being edited first.
3. The same part of the document was edited according to the Global English guidelines.
4. That part of the document was translated again.
5. The translations of 22 sentences were evaluated by professional translators.

As Table 1.1 shows, the translations of the Global English version of the document were significantly better than the translations of the unedited version. The percentage of sentences that were rated as either Excellent or Good increased from 27% to 68%. The percentage of sentences that were rated as either Medium or Poor decreased from 73% to 32%.

Table 1.1 Evaluations of Translations Produced by SYSTRAN, English-French

Translations of Unedited Sentences			Translations of Edited Sentences		
Rating	Number of Sentences	Percentage	Rating	Number of Sentences	Percentage
Excellent	0	0.00	Excellent	6	27.27
Good	6	27.27	Good	9	40.91
Medium	13	59.09	Medium	7	31.82
Poor	3	13.64	Poor	0	0.00

The sample size was admittedly very small. However, the results are consistent with results reported by Roturier (2006) and with what common sense tells anyone who has worked with computers: the quality of the output depends largely on the quality of the input.

More and more large companies are using machine translation successfully. They recognize that a certain amount of post-editing (corrections made by a human translator) is necessary in order to produce production-quality translations. But in many cases, production quality is not required. The goal might be simply to give readers the gist of a document's content.

When implemented intelligently and used selectively, machine-translation software reduces translation costs substantially. Equally important, machine translation can make it possible to provide rough translations of information that, for economic reasons, otherwise could not be translated at all.

For an excellent overview of machine translation, see Dillinger and Lommel (2004).

Translation Memory

Virtually all technical translators use computer-assisted translation tools. One of the main components of these tools is translation memory (TM)—a database that stores the source-language version and the target-language version of every sentence that is translated. When a new or updated document is processed by the software, any translation segments that are identical or similar to previously translated segments are presented to the translator. The translator then decides whether to reuse, modify, or disregard the previous translations.

8 *The Global English Style Guide: Writing Clear, Translatable Documentation for a Global Market*

Unnecessary inconsistencies make the use of translation memory less efficient. For example, suppose that a French translator translates the following sentence in the first edition of a software manual:

- ▶ Use the Group column to see if your tables are joined in more than one group.

Later, a translator who is using the TM database from that first edition encounters the same sentence in a new version of the document or in a related document. Instead of retranslating the sentence, the translator can insert the previous translation with the click of a mouse or with a keyboard shortcut.

Now suppose that the writer or editor who worked on the second edition of the manual had decided to modify that sentence as follows:

- ▶ Use the Group column to **determine whether** your tables are joined in more than one group.

In this scenario, the translation-memory software finds the translation of the original sentence in its database and presents that sentence and its translation to the translator as a fuzzy match. However, now the translator has to decide whether the previous translation is suitable or whether it needs to be modified. Obviously, that task is more cognitively demanding and more time-consuming than inserting the translation of an exact match. When the English sentence is ambiguous or difficult to understand, the task is especially time-consuming, and the decision process is subject to error.

When you multiply the unnecessary variations in a document by the number of languages that the document will be translated into, the cost of those variations becomes very significant. Therefore, many of the Global English guidelines are aimed at eliminating sources of unnecessary variation.

Controlled-Authoring Software

Learning to follow all of the Global English guidelines could be a daunting task—although one could argue that it is no more daunting than learning the guidelines that are in any other style guide. However, there is one technology that greatly facilitates the task of following not only the Global English guidelines, but many other style guidelines as well. That technology is commonly referred to as controlled-authoring software.³

³ The terms controlled-language checker and automated editing software are also used. However, the latter term sometimes refers to less sophisticated and non-customizable products that cannot provide adequate support for the Global English guidelines. See <http://www.globalenglishstyle.com> for a list of software vendors that license controlled-authoring software.

A controlled-authoring application parses texts and brings style errors, grammar errors, and terminology errors to the user's attention. One essential feature that distinguishes controlled-authoring software from other types of editing tools is that you can customize it. In collaboration with the software vendor, you can specify which grammar rules, style guidelines, and terminology restrictions you want the software to help authors follow.

You can also customize the rules to eliminate false alarms that are caused by idiosyncrasies in your documentation. Thus, the software is more accurate and reliable than off-the-shelf language checkers.

Many organizations use controlled-authoring software to ensure a high degree of language quality and consistency in their publications; to increase the productivity of content authors, editors, and translators; to help non-native authors produce better-quality English source texts; or for other business reasons.

Case Study

At SAS Institute, the implementation of controlled authoring was motivated partly by the need to standardize and control terminology. In recent years, SAS software products have become more integrated. SAS also began publishing documentation on the Web, with a consolidated index and full-text search. Terminology issues became more visible, both internally and to customers, than ever before.

The intensified pace of globalization also meant that SAS needed to find an efficient way of making its documentation more suitable for translation and easier for non-native speakers of English to understand. An earlier version of the Global English guidelines was developed for that reason and became an official part of the *SAS Style Guide for User Documentation*.

But even the best technical writers find it difficult to apply complex style guidelines or to consistently conform to lists of approved and deprecated terms. Deadlines and time pressures make it impractical for authors and editors to refer to style guides and glossaries frequently.

To emphasize the goal of helping authors communicate clearly and consistently, SAS used Assisted Writing and Editing (AWE) as the name of the project that encompassed the use of controlled-authoring software. After selecting a controlled-authoring product, SAS worked with the vendor to make the software as accurate as possible. For example, the software initially flagged the following sentence as an error because *the at* seemed to be an ungrammatical sequence of words:

- ▶ The remaining seven characters can include letters, digits, underscores, the dollar sign (\$), or **the at** sign (@).

That false alarm was eliminated by modifying the rule so that it ignores any occurrence of *the at* that is immediately followed by *sign*.

In the following sentence, the controlled-authoring software initially flagged *a HMDA* as an error and suggested *an HMDA* instead:

- ▶ To view a **HMDA** Edit Analysis Report, complete these steps:

But *HMDA* is pronounced as an acronym (HUM-dah), not as an initialism (H-M-D-A). Therefore, *a HMDA* was added to an exclusion list so that it would no longer be flagged as an error.

Controlled-authoring software gives authors immediate feedback on their own writing, teaching them to follow guidelines that they might otherwise have difficulty understanding or remembering. SAS has found that after an initial productivity hit, this training effect leads to the opposite: a significant productivity *increase*. Because authors fix grammar, spelling, style, and terminology issues early in the writing process, there are fewer corrections to be made late in the documentation cycle, when the pressure to deliver is greatest.

The software's consistent, objective feedback reduces unnecessary variation. SAS anticipates that the increased consistency in its documentation will make the use of translation memory more effective, and that consistent terminology and phrasing will make its documentation more usable for all audiences.

SAS is working closely with the vendor to further customize the software so that it will detect violations of more of the Global English guidelines. The software already detects violations of most of the other style guidelines and terminology restrictions in the *SAS Style Guide*.

For more details about implementations of controlled-authoring software, see Akis and Simpson (2002) and Kohl (2007).

Practical Considerations for Implementing Global English

Prioritize the Guidelines

Whether you use controlled-authoring software or not, you will probably want to focus on a subset of the Global English guidelines first. To help you decide which guidelines are most important for your circumstances, the heading for each major style guideline is followed by a Priority line that looks like this:

Priority: HT1, NN2, MT3

The following tables explain the acronyms and priority levels:

Acronym	Meaning	Priority Level	Meaning
HT	human translation	1	high priority
NN	non-native speakers	2	medium priority
MT	machine translation	3	low priority

For example, HT1 indicates that the guideline has high priority for documents that will be translated by human translators. NN2 indicates that the guideline has medium priority for untranslated documents that will be read in English by non-native speakers. And MT3 indicates that the guideline has low priority for documents that will be translated using machine-translation software. These priority values are based on the author’s subjective assessments and on feedback from translators and other localization professionals.

In Appendix B, “Prioritizing the Global English Guidelines,” the style guidelines are presented in tables that are sorted by the HT, NN, and MT values.

Build a Relationship with Your Localization Staff

If you don’t already know who manages the localization of your products and documentation, find out! Let them know that you are working toward making the localization process more efficient by improving the quality of your English documentation. Open a communication channel so that you will have someone to turn to when you have questions about whether a particular issue poses a problem for translation or localization. You, in turn, can point them to the right person if they have a question about the content of a particular document or product.

Always provide a glossary to your localization coordinator before the localization process begins. For more information about what to include in the glossary, see guideline 3.7.1, “Consider defining or explaining noun phrases.”

Eliminate Non-essential Information

In addition to following the Global English guidelines, be sure to consider other ways of reducing translation costs. One of the best ways is to reduce the volume of information to be translated. Content reduction can be done at the topic level, at the sentence level, or both.

Topic-Level Content Reduction

Many technical documents contain topics that are of interest to only a small percentage of readers. For example, as part of a “Downsizing Our Documentation” campaign at SAS Institute, a team of technical writers, software developers, and technical support analysts was able to remove 30% of the content of a 500-page technical reference manual. Many

topics in the document were of interest primarily to the technical support analysts and were therefore relocated to an internal Web site that is not translated.

According to one of the software developers, some of the information was there “to fill the term paper requirement.” That is, the corporate culture seemed to require that if you developed new functionality, the functionality had to be documented in the user manual, even if it was of interest to only a few customers. Obviously, a new feature that “requires” twelve pages of documentation seems more impressive than a new feature that “requires” only six. No one on the downsizing team was previously aware that the document was being translated into six languages at an average cost of \$.25 per word for each of those languages.

The technical writer was able to eliminate an additional 10% of the content by improving the organization of the document. He consolidated topics that were addressed in multiple places, and he was able to eliminate unnecessary introductions by using more-descriptive headings.

The difficulty with the topic-level approach is that it can require a considerable amount of time, effort, and coordination. Many organizations are not committed enough to the goal of reducing localization costs to assemble a team that has the right qualifications for deciding which content can be removed. Other priorities take precedence—especially since the division that pays for localization (and which would therefore reap the benefits of content reduction) is usually separate from the division that produces the documentation.

Sentence-Level Content Reduction

As the examples in Appendix A, “Examples of Content Reduction,” illustrate, even essential topics can usually be shortened by removing unnecessary sentences and by making remaining sentences more concise.

Unlike topic-level content reduction, sentence-level content reduction can be done by an individual author or editor, or by a team of authors and editors. The participation of subject-matter experts from other divisions is not required. The advantage of a team is that team members develop shared strategies for reducing content, which can then be applied to many documents.

With practice, it is possible to focus on the Global English guidelines and on sentence-level content reduction at almost the same time. Throughout this book, you are frequently encouraged to find a more concise way of expressing an idea instead of merely applying a Global English guideline.

For more information about content reduction, see Rushanan (2007) and Fenstermacher (2006).



Don't Eliminate Syntactic Cues

Even though you should always look for opportunities to be more concise, don't remove syntactic cues from your documents. Syntactic cues are function words, punctuation marks, and other language features that are optional in some contexts. For example, in the sentence below, the word *that* is a syntactic cue. It can be omitted without making the sentence ungrammatical, but its presence makes the sentence structure more explicit.

- ▶ Ensure **that** the power switch is turned off.

Many of the Global English guidelines encourage you to use syntactic cues in order to eliminate ambiguities and to improve the readability of your sentences. Therefore, syntactic cues should not be removed in order to reduce word counts.

Syntactic cues are discussed in detail in Chapter 6, "Using Syntactic Cues," and in Appendix D, "Improving Translatability and Readability with Syntactic Cues."

Insert Explanations for Translators

As you will see when you read other parts of this book, sometimes a clause or sentence is ambiguous and there is no practical way to make it unambiguous. It is best to prepare for that situation by having a standard way of inserting explanations into your text for your translators' benefit.

Here is an example. In the following sentence, it is not entirely clear whether the relative clause, *that contains the data source*, modifies *Folders tree* or *location*:

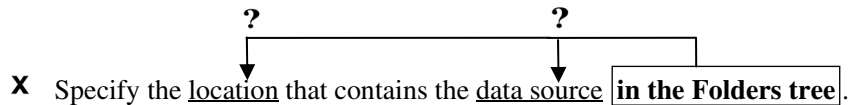
- ? ?
- ↓ ↓
- ▶ Specify the location in the Folders tree that contains the data source.

In other words, does the location contain the data source, or does the Folders tree contain the data source?

Suppose the author knows that the relative clause modifies *location*. Usually it is better to place a relative clause as close as possible to what it is modifying:

- ▶ Specify the location **that contains the data source** in the Folders tree.

But the above revision causes a different ambiguity. Now a translator might misinterpret the prepositional phrase, *in the Folders tree*, as modifying *data source*.



In fact, it, too, modifies *location*. That is, the location contains the data source, and the name of the location is displayed graphically in a tree-like, hierarchical list of folders.

For XML documents, the Internationalization Tag Set (ITS), Version 1.0, includes a Localization Note data category that is used for this purpose. (See <http://www.w3.org/International/its>.) You can use the <locNote> tag to include localization notes for your translators, as in the following example:

- ▶ Specify the location in the Folders tree that contains the data source.
`<locNote>"that contains the data source" modifies
"location"</locNote>`

Whatever publishing tool you use, it should provide some way of inserting comments that are visible only when another user of those tools specifies a particular setting or option. Thus, you can provide these explanations to translators without including the explanations in your deliverables.

For more examples of contexts in which you might need to provide explanations to translators, see guidelines 4.2.5, 4.6.7, 5.1.1, and 5.1.2.

Frequently Asked Questions about Global English

What is the relationship between Global English and controlled English?

First, controlled English is not a single entity. The term describes any of several attempts to define a subset of the English language that is simpler and clearer than unrestricted English. Most versions of controlled English specify which grammatical structures are allowed and which terms are allowed, as well as how those terms may be used. In early forms of controlled English, terminology was often restricted to a core vocabulary (in some cases, as few as 800–1000 terms), supplemented by technical terms that are necessary for a particular subject area or product.

Global English could be regarded as a loosely controlled language, yet it was developed using almost the opposite approach. In the development of Global English, the emphasis has been on identifying grammatical structures and terms that should be *avoided*, rather than on cataloging all of the grammatical structures and terms that are *allowed*. In other words, anything that is not specifically prohibited or cautioned against is allowed.

When texts conform to the guidelines and terminology restrictions of the more restrictive forms of controlled English, the style and rhythm of those texts differs noticeably from the style and rhythm of unrestricted technical English. By contrast, most readers don't notice anything different about the style and rhythm of texts that conform to the Global English guidelines.

Early versions of controlled English, such as the Kodak International Service Language (KISL), were developed as alternatives to translation. By severely limiting the range of grammatical structures and vocabulary that are allowed, KISL makes technical documents understandable even to readers who have very limited proficiency in English. Kodak found that it is much less expensive to teach service technicians all over the world a limited amount of English than to translate service manuals into 40 or more languages.

Global English is an alternative to translation only if the non-native speakers in your audience are reasonably proficient in English. Global English can make the difference between documents that those non-native speakers can read easily and documents that are too difficult for that audience to comprehend.

If you are writing for readers who have limited proficiency in English, then consider using a form of controlled English. However, keep in mind that the amount of effort and knowledge that is required for developing and implementing controlled English is considerable. Consult the Bibliography section of this book for sources of more information about controlled English.

Do the Global English guidelines make all sentences clear and easy to translate?

As Farrington (1996) said, "Simplified English will not compensate for a lack of writing skills." He was referring to AECMA Simplified English, the forerunner to ASD Simplified Technical English,⁴ but the same could be said of Global English. Authors who follow the Global English guidelines still have to have good writing skills in order to produce high-quality prose.

Does following these guidelines lead to an increase in word counts?

According to Bernth (1998b), "In order to cut down on ambiguity, it is nearly always necessary to be somewhat verbose." However, the guidelines in this book include frequent reminders to search for opportunities to be more concise. Authors who pay attention to those reminders are able to eliminate unnecessary text and verbiage while they are applying the Global English guidelines.

Even if your word counts increase slightly, the benefits of increased clarity, readability, and consistency outweigh the additional per-word cost of translation.

⁴ ASD Simplified Technical English is a form of controlled English that was initially developed for use in the aerospace industry. It is now used in other industries as well.

Typographical Conventions

This book contains hundreds of example sentences. The following symbols indicate what each example is intended to illustrate or represent:

Symbol	Meaning
X	An example that, from a Global English perspective, is incorrect or undesirable.
✓	A revised example that conforms to the Global English guideline that is being presented.
✓ ₄	A revised example that conforms to the Global English guideline that is being presented and that is stylistically better than a previous revision.
?	A revised example that might be acceptable to some people but not to others, or that might be technically incorrect.
⊗	An example that conforms to Global English guidelines but that is either technically incorrect or stylistically unacceptable.
▶	An example that illustrates a point of grammar or style rather than illustrating compliance or non-compliance with a Global English guideline.
T	A translation of an example sentence.
