

The Top Ten Global English Guidelines

by John R. Kohl

The correct bibliographic citation for this document is as follows: Kohl, John R. 2005. *The Top Ten Global English Guidelines*, Cary NC: SAS Institute Inc.

The Top Ten Global English Guidelines

Copyright © 2005, SAS Institute Inc., Cary, NC, USA.

All rights reserved. Produced in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

U.S. Government Restricted Rights Notice: Use, duplication, or disclosure of the software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

April 2005

SAS Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at support.sas.com/pubs or call 1-800-727-3228.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

Contents

Introduction.....	1
The Cardinal Rule of Global English.....	2
1. Use short sentences.....	2
2. Use complete sentences	4
3. Untangle long noun phrases.....	6
4. Expand -ED verbs whenever possible	6
5. Always revise -ING verbs that follow nouns.....	7
6. Use "that" liberally.....	8
7. Choose simple, precise words that have a limited range of meanings.....	8
8. Standardize your terminology and phrasing	9
9. Don't use slang, idioms, colloquialisms, or figurative language	11
10. Clarify which parts of a sentence are being joined by "and" or by "or."	12

Introduction

This document provides a brief, simplified version of 10 of the approximately 80 Global English guidelines that have been formulated over the past 12 years at SAS Institute. Global English refers to two general types of guidelines that can make written texts easier to translate or easier for non-native speakers of English to read and comprehend:

- Sentence-level guidelines, which pertain to the clarity, accuracy, readability, or translatability of sentences or parts of sentences (for example, various types of phrases and clauses).
- Terminology guidelines, which pertain to individual terms and concepts (including a few higher-level abstractions such as humor and satire that are sometimes expressed in individual words) and to short phrases.

The Importance of Customization

This document was produced as an internal document, and the guidelines described here were judged appropriate for SAS software documentation. The document is being made available in its present form to individuals outside of SAS who are interested in Global English. However, readers of this document should decide for themselves whether all or some of these guidelines are useful and appropriate in other contexts.

Where to Find the Complete Global English Guidelines

The complete version of the Global English guidelines, with additional information about Controlled English software, machine translation, and other language technologies, is slated for publication sometime in 2006. The working title of the book is *Global English: Writing Clear, Translatable Documentation for a Global Market*. The book will probably also include access to online lessons and to other online materials. You can check on the status of the book at <http://support.sas.com/companionsites>.

Typographical Conventions

Next to the example sentences, phrases, and terms, this document uses the following symbols:

Symbol	Meaning
×	An example that is incorrect or undesirable.
?	A revised example that might be acceptable to some people but not to others, or that might be technically incorrect.
☑	A revised example that conforms to the Global English guideline that is being presented.
+	A revised example that conforms to the Global English guideline that is being presented and that is stylistically better than a previous revision.

About the Examples

Because SAS is a software company, most of the example sentences and paragraphs in this document are from software documentation.

The Cardinal Rule of Global English

At SAS, native speakers of English still constitute the majority of the audience for most of our documentation. Therefore, although we take the needs of non-native speakers and translators into account, we emphasize the Cardinal Rule of Global English:

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

At the same time, we hasten to add the following corollary:

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

In other words, if following one of these guidelines results in a sentence that sounds stilted or unnatural, then try to find a different way to improve the sentence.

1. Use short sentences.

In conceptual information, limit your sentences to 25 words. For task-based or procedural information, limit your sentences to 20 words. Of course, these are general guidelines, not absolute limits for every single sentence.

- ✗ With design-time controls, you control the look and feel of your Web pages in a WYSIWYG editor environment, and at the same time use all the functionality of SAS/IntrNet software in your Web pages. (35 words)

- ☑ With design-time controls, you control the look and feel of your Web pages in a WYSIWYG editor environment. **In addition, you can** use all the functionality of SAS/IntrNet software in your Web pages. (19 + 15 = 34 words)
- ✗ You can use this application to review survey information that was entered previously, as well as to update or add survey responses for new batches of ice cream. (28 words)
- ☑ You can use this application to review survey information that was entered previously. You can also add survey responses for new batches of ice cream. (13 words + 12 words = 25 words)
- ✗ Remote Object Class Factory: a library of Java classes that enable applications developers to automatically encapsulate SAS/AF software models so that a client application can use them without requiring any specialized code. (28 words in the definition)
- ☑ Remote Object Class Factory: a library of Java classes that enable applications developers to automatically encapsulate SAS/AF software models. A client application can then use the models without requiring any specialized code. (15 + 13 = 28 words in the definition)

Long sentences aren't always a problem.

A long sentence is not necessarily problematic if it consists of two or more clauses, each of which is clearly written. But you should examine all long sentences to see whether they could easily be divided with no loss of coherence. Here's a long sentence that is nearly impossible to comprehend:

- ✗ If you name a member type or several member types in the PROC DATASETS statement, in most subsequent statements in the PROC DATASETS step (the CONTENTS and COPY statements are exceptions to this rule), you can name only a subset of the list of member types included in the PROC DATASETS statement. (52 words)

In the following revised version of the sentence, notice that the first sentence is 31 words long—6 words over our suggested limit for conceptual information. However, that long sentence consists of two clauses, each of which is clearly written.

- ☑ If you specify one or more member types in the PROC DATASETS statement (13 words), then you can usually specify only a subset of those member types in subsequent statements in the procedure (18 words). The CONTENTS and COPY statements are exceptions to this rule. (10 words)
13 + 18 + 10 = 41 words total

Also notice that the original sentence was not merely divided into two sentences: it was revised considerably and was made more concise. It's always good to think about different ways of improving a problematic sentence rather than blindly following the Global English guidelines and applying the first remedy that comes to mind.

Sentences should sometimes be divided even if they don't exceed the limit.

If there is an easy way to divide a sentence that is somewhat shorter than the recommended limit, you should consider doing so. In the following glossary definition, the relative clause (*which enables...*) can easily be made into a separate sentence.

- ✗ CGI-based technology: a technology that is based on the Common Gateway Interface (CGI) standard, **which** enables external applications to communicate with information servers such as Web servers. (25 words in the definition)
- ☑ CGI-based technology: a technology that is based on the Common Gateway Interface (CGI) standard. **CGI** enables external applications to communicate with information servers such as Web servers. (12 + 13 = 25 words in the definition)

2. Use complete sentences.

Incomplete or interrupted sentences can cause extra work for translators, because the order of nouns, verbs, prepositional phrases, and other sentence constituents is different in other languages. In addition, incomplete or interrupted sentences can cause translation software to produce garbled results.

Use complete sentences in introductions to lists.

Example 1

- ✗ In addition to invoking, managing, and scrolling windows, the windowing environment can
 - customize windows
 - manage libraries and files
 - search text.
- ☑ In addition to invoking, managing, and scrolling windows, the windowing environment can be used as follows:
 - to customize windows
 - to manage libraries and files
 - to search text.

Example 2

- ✗ In order to use formatting macros, you must be familiar with Base SAS software, including
 - using the SAS windowing environment
 - creating and submitting simple SAS programs
 - using the SAS macro language.
- ☑ In order to use formatting macros, you must know how to do the following:
 - use the SAS windowing environment
 - create and submit simple SAS programs
 - use the SAS macro language.

Don't interrupt sentences with programming code.

Example 1

- ✗ To automatically define a libref each time SAS starts, add


```
libname _saswa <location-of-your-Knowledge-Base>;
```

 to your autoexec.sas file.
- ☑ To automatically define a libref each time SAS starts, add the following statement to your autoexec.sas file:


```
libname _saswa <location-of-your-Knowledge-Base>;
```

Example 2

- ✗ The following code


```
url symget('_URL');
line='<A HREF=*'||trim(left(url))||
put line;
```

 produces


```
<A HREF="/cgi-bin/broker ?_service=default&_program>
```

- ☒ Consider the following code:

```
url symget('_URL');  
line='<A HREF=*'|trim(left(url))|'  
put line;
```

This code produces the following output:

```
<A HREF="/cgi-bin/broker ?_service=default&_program>
```

3. Untangle long noun phrases.

As this example illustrates, it is often difficult for translators, and even for users, to determine what a long noun phrase means.

- ✗ The **default column pointer location** is column 1.

Does “default column pointer location” mean “the location of the default column pointer” or “the pointer location of the default column” or “the default location of the column pointer”? Only a subject-matter expert could know for sure that the last interpretation is correct.

- ✗ The **location of the default column pointer** is column 1. (incorrect interpretation)
- ✗ The **pointer location of the default column** is column 1. (incorrect interpretation)
- ☒ The **default location of the column pointer** is column 1.

In your revisions, omit any unnecessary words.

In the following heading, the product name was superfluous. The entire document was about SAS/ABC software, and it was clear from the context that we were discussing SAS/ABC table cells. Therefore, the last revision is the only satisfactory solution.

- ✗ Editing **SAS/ABC Table Cell Properties**
- ✗ Editing **the Properties of SAS/ABC Table Cells**
- ☒ Editing **the Properties of Table Cells**

4. Expand -ED verbs whenever possible.

Note: The rationales for guidelines 4, 5, 6, and 10 are explained in Kohl, John R. “Improving Translatability and Readability with Syntactic Cues.” *Technical Communication* 46.2 (May 1999): 149-166.

If the simplest revision sounds unnatural, then try to find a better solution.

This example shows that there are different ways of revising -ED verbs. Many people would consider the second revision to be better stylistically.

- ✗ With the information **provided** in WebHound, you can assess the present level of service.
- ☑ With the information **that is provided** in WebHound, you can assess the present level of service.
- + With the information **that** WebHound **provides**, you can assess the present level of service.

A drastic revision might be the best solution.

Instead of expanding every -ED verb into a relative clause, consider other alternatives—especially when multiple -ED verbs are close together in a sentence.

- ✗ In order to display labels **specified** in a LABEL statement **used** with PROC PRINT, you must specify the LABEL option in the PROC PRINT statement.
- ✗ In order to display labels **that are specified** in a LABEL statement **that is used** with PROC PRINT, ... (unnatural)
- ☑ In order to display labels with PROC PRINT, you must do two things:
 - Specify the labels in the LABEL statement.
 - Specify the LABEL option in the PROC PRINT statement.

If you can't think of an acceptable revision, then leave the sentence alone.

Maybe your editor will find an elegant solution!

5. Always revise -ING verbs that follow nouns.

-ING verbs are problematic in other contexts as well, but when they follow nouns, you should always revise them.

There are many ways of expanding, revising, or eliminating -ING verbs that follow nouns.

- ✗ To explain the capabilities **being** defined, you can insert comment lines **following** each block of code.
- ☑ To explain the capabilities **that are being** defined, you can insert comment lines **after** each block of code.
- ☑ To explain the capabilities **that you are defining**, you can insert comment lines **after** each block of code.

When there is more than one noun + -ING verb in a sentence, it does not necessarily sound unnatural to expand both of them into relative clauses. However, consider ways of avoiding such repetition.

- ✗ A database server **running** on a mainframe might be responsible for looking up records for many clients **running** on individual users' workstations.
- ☑ A database server **that is running** on a mainframe might be responsible for looking up records for many clients **that are running** on individual users' workstations.
- ☑ **If** the database server **is running** on a mainframe, it might be responsible for looking up records for many clients **that are running** on individual users' workstations.

6. Use "that" liberally.

The following verbs often take "that" complements: *assume*, *ensure*, *indicate*, *mean*, *note*, *require*, *specify*, and *verify*.

- ✗ Check for syntax errors and ensure the macro name and the member name are the same.
- ☑ Check for syntax errors and ensure **that** the macro name and the member name are the same.

As usual, there are often alternatives to simply inserting "that."

- ✗ The rules you define might depend on the type of data you are importing.
- ☑ The rules **that** you define might depend on the type of data **that** you are importing.
- + The rules **that** you define might depend on **which** type of data you are importing. (sounds much better)

7. Choose simple, precise words that have a limited range of meanings.

Most words have multiple meanings, so try to become more aware of how particular words might be interpreted by translation software. If you are using one of the more unusual meanings of a word, perhaps there is a different word that you could use instead—one that has only the one particular meaning that you are trying to convey.

You should also avoid unusual words that non-native speakers of English are not likely to be familiar with.

Examples

"Parenthesized" is in the dictionary, but it's quite unusual, and it might be puzzling to some non-native speakers:

- ✗ %%EVAL must be followed by a **parenthesized** expression.
- ☑ %%EVAL must be followed by an expression **that is enclosed in parentheses**.

"Uncondense" and "subdiagram" are not in Webster's dictionary. "Uncondense" can easily be replaced by "expand." "Subdiagram," on the other hand, is a legitimate technical term. When in doubt, consult your editor.

✗ You cannot **uncondense** the nodes in a **subdiagram**.

☑ You cannot **expand** the nodes in a subdiagram.

"Roughly" has quite a different meaning in other contexts, as in "The guard treated the prisoner roughly." By contrast, "approximately" has only one common meaning.

✗ Encryption requires **roughly** the same amount of CPU resources as compression.

☑ Encryption requires **approximately** the same amount of CPU resources as compression.

At SAS, we use "might" instead of "may" if we're writing about a possibility or an option rather than about permission.

✗ You **may** need to change the value of the MVARSIZE system option.

☑ You **might** need to change the value of the MVARSIZE system option.

Sometimes "can" can be used instead. However, there might be some contexts in which nothing but "may" sounds appropriate.

Don't go to extremes.

It is not always necessary to use the most literal or most common meaning of a particular word. For example, here are some common meanings of two words that occur frequently in SAS documentation:

illustrate—to provide with visual features (to *illustrate* a book).

reflect—to prevent passage of and cause to change direction (a mirror *reflects* light); to think quietly and calmly.

But within SAS documentation, these terms are virtually always used to convey other meanings, as in the following sentences.

☑ Figure 20.1 **illustrates** the process of creating and analyzing variance reports.

☑ The term *input buffer* refers to a logical concept and does not necessarily **reflect** the physical storage of data.

Human translators would surely understand these meanings, and few non-native speakers would be puzzled by them. Even if we used translation software, we could simply program the software to always translate these words according to how they are typically used in our documentation.

8. Standardize your terminology and phrasing.

Avoid using different words to convey the same meaning.

The word "adjust" doesn't have a lot of different meanings. However, in SAS documentation, as well as in general discourse, "change" is much more common, and in the following sentence, it has essentially the same meaning.

- ✗ **Adjust** the value of the MSYMTABMAX system option if necessary.
- ☑ **Change** the value of the MSYMTABMAX system option if necessary.

Similarly, there is nothing inherently wrong with "supply." But we use "specify" more frequently in similar contexts, so it would be good to standardize on that term.

- ✗ The description that you **supplied** for the ZYX option was truncated.
- ☑ The description that you **specified** for the ZYX option was truncated.

Avoid using one word to convey different meanings.

In this example, the first instance of the verb "enable" means "to give someone or something the capability to do something." (**Note:** At SAS, we never use "allow" to convey this meaning. We use "allow" only to mean "to give permission.") The second instance of "enable" means "to activate." There is no justification for using "enable" to convey that second meaning. It is easy to find alternatives.

- ✗ You can **enable** users to create their own bookmarks. ... In order to **enable** run-time bookmarks, two process must occur.
- ☑ ... In order **to give users this capability**, two processes must occur.
- ☑ ... In order **for users to create their own bookmarks**, you must do the following:
- ☑ ... **To enable** users to create their own bookmarks, you must do two things:

Standardize your phrasing.

Try to be consistent at least within the same document, so that you aren't saying the same thing in many different ways. Of the following phrasings, the ones in bold are the ones that we've tried to standardize on at SAS.

- ☑ **Enclose the attribute value in parentheses.**
- ✗ Surround the attribute value with parentheses.
- ✗ Put parentheses around the attribute value.
- ✗ Place parentheses around the attribute value.
- ✗ Specify the attribute value, enclosed in parentheses.
- ☑ **The attribute value must be enclosed in parentheses.**
- ✗ The attribute value must be surrounded by parentheses.

- ✗ The attribute value must be within parentheses.
- ✗ Parenthesize the attribute value.

9. Don't use slang, idioms, colloquialisms, or figurative language.

Figurative language and idioms

- ✗ Clarke's definition is **on target**, but it might not **capture the full flavor of** the important distinctions in approaches to usability engineering.
- ☑ Clarke's definition is **accurate**, but it does not **account for** [or **fully explain** or ??] important distinctions in approaches to usability engineering.
- ✗ We have created this site to **keep you posted** about the newest enhancements to our software.
- ☑ We have created this site to **keep you informed** about the newest enhancements to our software. ("to inform you of" sounds too stilted)

Note: Sometimes there is no acceptable alternative to an idiomatic expression.

Idiomatic multi-word verbs

- ✗ Use the MEMTYPE= option in the following statements to **cut down on** the member types that are available for processing.
- ☑ Use the MEMTYPE= option in the following statements to **limit** the member types that are available for processing.
- ✗ The details of how user exits are implemented are **left up to** each host to determine.
- ☑ User exits are implemented differently on each host.
- ✗ This method subsets the data and **finds out** which variable in the hierarchy to display next.
- ☑ This method subsets the data and **determines** which variable in the hierarchy to display next.
- ✗ Positive numbers **stand for** the top n values; negative numbers **stand for** the bottom n values.
- ☑ Positive numbers **represent** the top n values; negative numbers **represent** the bottom n values.

Note: For some idiomatic multi-word verbs, there might not be a good substitute.

Jargon and colloquialisms

Jargon and colloquialisms are uncommon in technical documents, but here are some examples from other sources:

- ✗ **bleeding-edge** technology
- ✗ a **rule of thumb**
- ✗ a **quick-and-dirty** solution
- ✗ The **tricky thing** about this Java servlet is the rewriting of the URLs.
- ✗ Machine translation **boils down to** the ability of programmers to take the knowledge of computational linguistics and express it in programming.
- ✗ A Java servlet is **command central**.

10. Clarify which parts of a sentence are being joined by "and" or by "or."

The following syntactic cues make it clear which parts of the example sentences are being conjoined.

A preposition

- ✗ SAS/DQ enables you to standardize your data **by building** schemes from your data **and applying** those schemes to your data.
- ☑ SAS/DQ enables you to standardize your data **by building** schemes from your data **and by applying** those schemes to your data.

The infinitive marker "to"

- ✗ You cannot use the DATASETS procedure **to change** the values of observations, **change or delete** variables, **or change** the type or length of variables.
- ☑ You cannot use the DATASETS procedure **to change** the values of observations, **to change or delete** variables, **or to change** the type or length of variables.

"That"

- ✗ Typically, the message indicates **that** a variable's type has been incorrectly identified **or** the raw data file contains some invalid data values.
- ☑ Typically, the message indicates **that** a variable's type has been incorrectly identified **or that** the raw data file contains some invalid data values.

An auxiliary verb

- ✗ Because the expression produces a numeric value, TOTAL **is defined** as a numeric variable **and assigned** a default length of 8.
- ☑ Because the expression produces a numeric value, TOTAL **is defined** as a numeric variable **and is assigned** a default length of 8.

A modal verb

- ✗ Character values **must be enclosed** in quotes **and be** in the same case as in the data set.
- ☑ Character values **must be enclosed** in quotes **and must be** in the same case as in the data set.

Don't go to extremes.

You don't need to insert a syntactic cue if the parts of the sentence that "and" or "or" is joining are very short and close together, as in this example.

- ✗ The text substitution that this activity produces is complete before the program text **is compiled and executed**.

It would sound very unnatural to say "is compiled and **is** executed." (Remember the Cardinal Rule of Global English!)

Be careful not to accidentally distort your meaning.

If you insert the wrong syntactic cue, you might make the sentence technically incorrect. In the following example, which revision is correct? If you as the writer didn't know, then you would have to research the matter to find out.

- ✗ The values for SALARY begin in column 19, so use the pointer control **to** point to column 19 and name the variable.
- ? The values for SALARY begin in column 19, so use the pointer control **to** point to column 19 and **to** name the variable.
- ? The values for SALARY begin in column 19, so use the pointer control **to** point to column 19, and **then** name the variable.

Note: At SAS, if an editor inserts syntactic cues, it is the writer's responsibility to review the edits carefully to ensure that the editor resolved all of the ambiguities correctly.

