



Meditations for a Tranquil Technical Translator

or

How *Human* Translators, not Computer Scientists, Will Choose
the Standards for Controlled English

What are Controlled Languages?

“Controlled Natural Languages are subsets of natural languages whose grammars and dictionaries have been restricted in order to reduce or eliminate both ambiguity and complexity. Traditionally, controlled languages fall into two major categories: those that improve readability for human readers, particularly non-native speakers, and those that improve computational processing of the text.... STE is the most widely used CL in the world today.”

[Language Technology Group (LTG) of the Human Communication Research Centre (HCRC) at Macquarie University, Sydney]

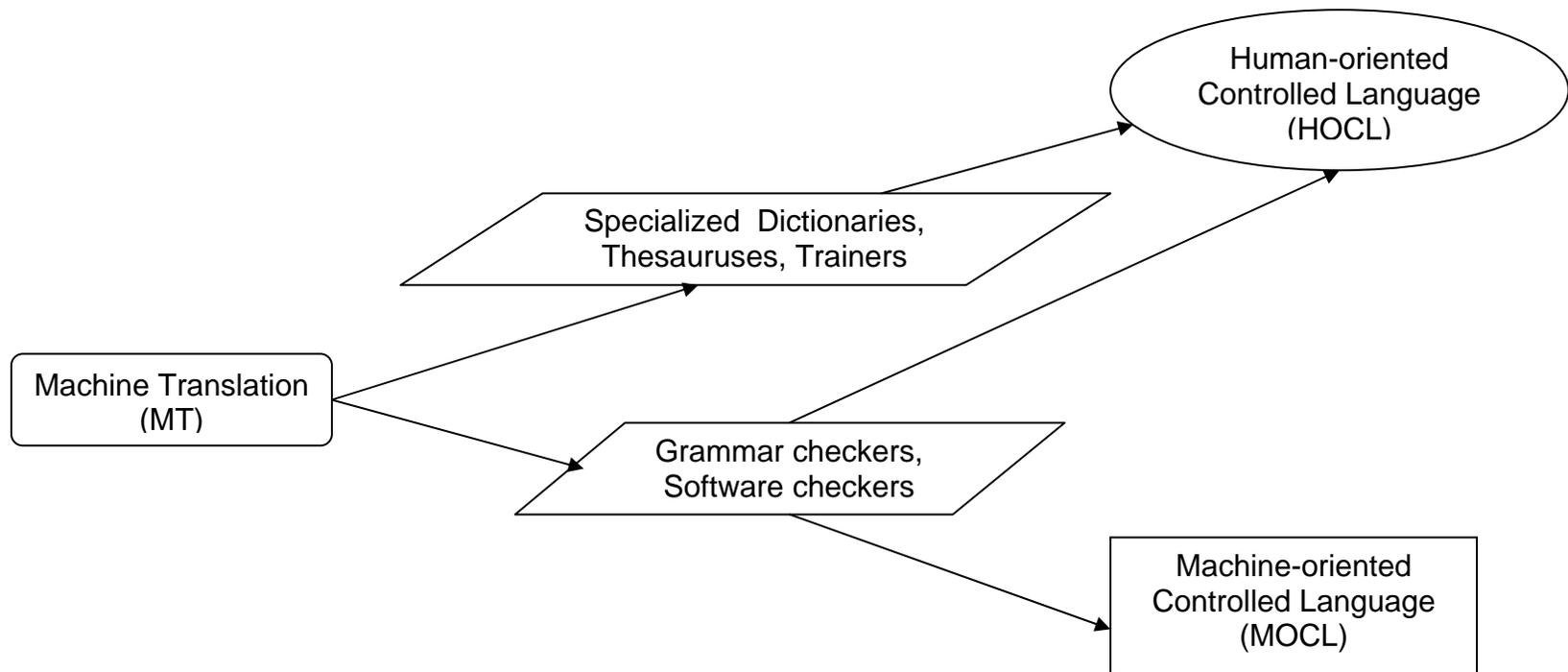
Four studies supporting the assertion that CL improves readability and comprehensibility, actually only study STE.

[p 85, Nyberg]

Many of its rules are recommendations found in technical writing textbooks. For example, SE requires writers to:

- Use the active voice.
- Use articles wherever possible.
- Use simple verb tenses.
- Use language consistently.
- Avoid lengthy compound words.
- Use relatively short sentences.

"Controlled Natural Languages are subsets of natural languages whose grammars and dictionaries have been restricted in order to reduce or eliminate both ambiguity and complexity. Traditionally, controlled languages fall into two major categories: those that improve readability for human readers, particularly non-native speakers, and those that improve computational processing of the text."



Controlled Englishes:

1. Basic English (1930s-40s)
2. International Language of Service and Maintenance (ILSAM)
 - a. Simplified Technical English or ASD-STE100 specification, formerly AECMA Simplified English (*Aerospace and Defense Industries Association of Europe*)*
 - b. IBM
 - c. Rank Xerox
 - d. Ericsson Telecommunications
3. Caterpillar Functional English 1970s/ Caterpillar Technical English
4. KANT/ Caterpillar Technical English
5. Six Sigma English
6. US government's plain English
7. ACE (Attempto Controlled English)*
8. Alcatel COGRAM*
9. IBM Easy English*
10. GM CASL*
11. OCE's Controlled English*
12. Sun's Controlled English*
13. Avaya's Controlled English*
14. PACE (Perkins Approved Clear English)
15. Common Logic Controlled English (CLCE)
16. Diebold Controlled English
17. COBOL
18. CASE (Case's Clear & Simple English)
19. Airspeak, Seaspeak, Policespeak (Cambridge University)
20. Douglas Aircraft Company dictionary, 1979

* An Analysis of Several Controlled English Rule Sets

Only one CL rule is common to all Rule Sets:

Keep (procedural) sentences as short as possible (20 words maximum)
(SE Rule 5.1)

There are seven rules shared by at least four of the CLs:

Use approved words from dictionary

Make instructions as specific as possible

Do not make noun clusters longer than three nouns

When appropriate, use an article or demonstrative adjective before a noun

Avoid the gerund

Use only active voice

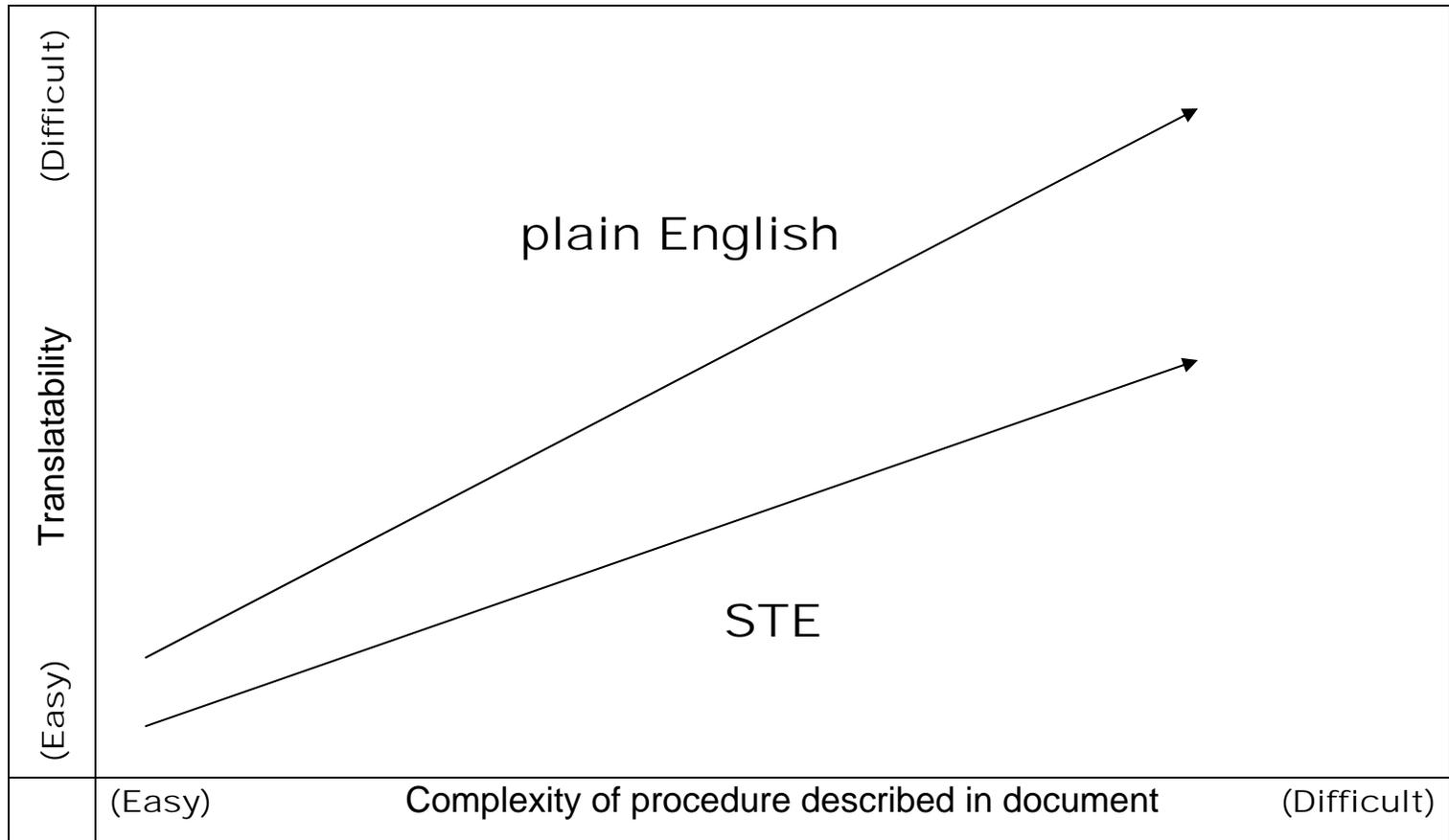
Do not omit relative pronouns such as “who”, “which” or “that”

[Source: <http://www.ctts.dcu.ie/O'Brien.ppt>]

Pluses of Using CLs

- Improved comprehensibility
- Easier post-processing and re-use of documents
- A measurable index of document quality
- Lower project management costs
- Lower translation costs
- Reduces revisions and costs at editorial level

Translatability of STE Relative to plain English



Up to 20% increase in time to prepare a source document in conformance with a Controlled English specification [Goyvaerts, 1996]

Vs.

Between 50% to 70% savings on translation costs [Pym, 1993]

Drawbacks

Conformance problem – Organizational patterns of language use and feedback are as important as a list of definitions, yet none of the CLs provide rules exhaustive enough to define the language it covers. Enforcement of conformance.

Authoring problem – The easier it is for the reader or translator, the harder it is for the author. Problem-reporting process.

Evaluation problem - Some of the elements in the definition of the CL rules, or in the list of words to avoid, impose burdens on the writer that are disproportionate to the benefits for the reader. This is not detected by merely evaluating a finished translation.

Portability and Customization problems – The goal of being stylistically uniform ignores the need to differentiate procedural manuals from maintenance manuals. A CL may become repetitive in the interests of clarity, especially an MOCL.

Sample conformance problem:

Consider the first vague rule, “Keep sentences short” and the second one which contradicts it “Avoid strings of nouns.” Adherence to the latter would suggest rewriting the seven-word phrase in (a) as the 11-word phrase (b), which clashes with the former writing rules.

- a. the nose landing gear uplock attachment bolt
- b. the bolt that attaches the uplock to the nose landing gear

Sample conformance success:

manual power tool - original text:

Gain access to blade. After removing old blade, new blade may be fitted by proceeding in reverse order, using gloves to avoid injuries by teeth of blade.

Before you attempt any of the above, the power should have been switched off.

manual power tool - text rewritten in Simplified Technical English:

Make sure that the on/off switch is in the "off" position.

Remove the blade cover from the machine.

Warning: wear gloves when you touch the blade.

Remove the old blade. Install the new blade. See figure A.

Install the blade cover.

Proscriptive, prescriptive CL checkers

- Proscriptive
 - Heuristic patterns or templates
 - Less work
 - More likely to give inappropriate feedback
- Prescriptive
 - Definition of each and every linguistic structure that is allowable in the CL
 - Labor-intensive
 - Fewer false alarms

With either CL checker, human judgement is necessary:

- There are likely to be some sentence structures that are overlooked in the original language definition but still considered necessary; hence the grammar rules must undergo extension and tuning during initial use.
- If the words are not near-synonyms, automatic correction is even harder, or even impossible. Resolving such situations usually requires human judgement.

[p. 80 "Controlled Language for Authoring and Translation,": Nyberg, Eric, Mitamura, Teruka and Huijsen Willem-Olaf, 1998]

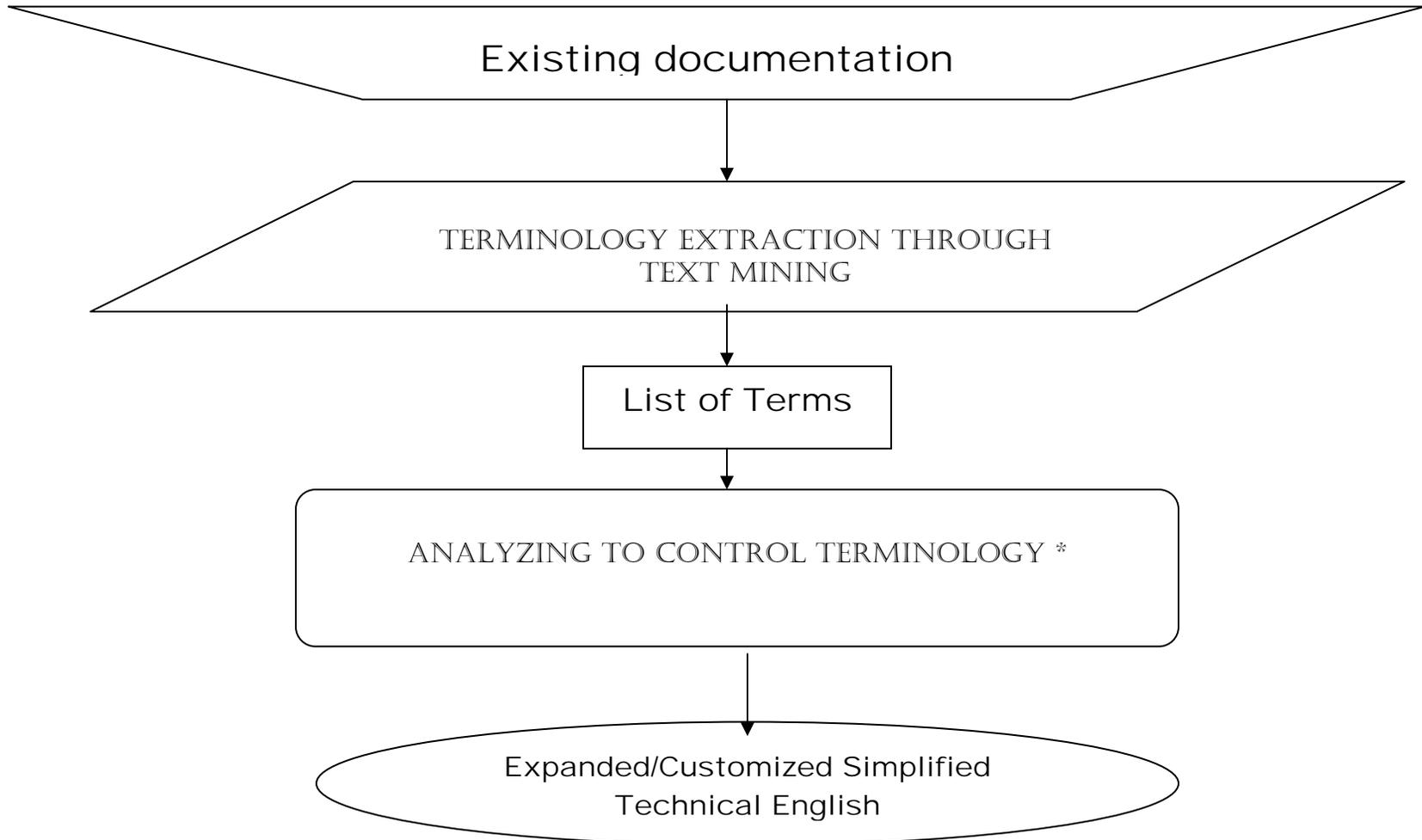
Text Grade Level Check

Use the Flesch-Kincaid Grade Level tool in Word. In Microsoft Word, on the toolbar, click on Tools. On the Tools menu, click on Options. On the tabs at the top of the box, click on Spelling and Grammar. Make sure both of the following have check marks:

Check grammar with spelling
Show readability statistics

This command instructs Microsoft Word to go through your document and do both a Spell Check and a Grammar Check, and to give you Readability Statistics which will give you data on Counts, Averages, and Readability. Under Readability, look for the Flesch-Kincaid Grade Level.

The Language Funnel



Rules for Expanding STE Vocabulary

- Choose the words from:
 - Approved words in the Dictionary
 - Words that qualify as Technical Names (Refer to Rule 1.5)
 - Words that qualify as Technical Verbs (Refer to Rule 1.10).
- Use a Technical Name only as a noun or an adjective, not as a verb.
- Some unapproved words are used to complete Technical Names. Do not use these unapproved words unless they are part of a Technical Name.
- Do not use different Technical Names for the same thing.
- Once you choose the words to describe something, continue to use these same words (particularly Technical Names).
- Use Technical Verbs only as verbs, not as nouns (unless the noun form qualifies as a Technical Name).
- You can use the past participle of the Technical Verb as an adjective.



Correction as Translation: Two Thoughts

1. The heuristic approach uses pattern substitution methods to correct violations. A more principled approach is known as correction as translation. It formulates correction as a translation problem, so that MT technology can be applied.
2. According to CLAW proceedings at the Kyoto meeting [Shirai, 1998], automatic rewriting rules are often used for a Japanese-English MT system because the syntax of the two languages is very different and it is useful to transform the input sentences before running them through MT. An experiment in automatic rewriting shows that the quality of Japanese-English MT is improved by 20% when rewriting rules are applied.

Author Memory versus Translation Memory

Know all the steps of your authoring and translation processes very well before applying any technology to them.

Do not lose focus on MT post-editing (PE) environment so that it does not attain the same quality level as the CL authoring pre-editing environment.

Text mining can introduce problems of synonymy into an existing CL.

Save your client/employer from themselves -
Read product reviews before you sign on for Machine Translation!

The following journals may be useful:

IDC Reports

TWS Language Technology Research Center

ELDA

MLC&T

International Journal of Language & Documentation

Language International

STE Lessons for IJET-16

1. Use approved words from the Dictionary only as the part of speech given.

Example: "Test" is approved as a noun but not as a verb.

WRITE: Do the leak test for the system.

or

Do a test for leaks in the system.

NOT: Test the system for leaks.

Example: "Close" is a verb (and not an adverb).

WRITE: Do not go near the landing gear if ...

NOT: Do not go close to the landing gear if ...

2. When appropriate, use an article (the, a, an) or a demonstrative adjective (this, these) before a noun.

Example:

WRITE: Lift up the assembly and put it in a box.

NOT: Lift up assembly and put in box.

Articles are not necessary before all nouns in a series or before mass nouns used in general statements.

Examples:

Put the packings, gaskets, and seals in a safe location.

Solvents used in these repairs can cause damage to paint.

3. To change a passive construction to the active, you can use these methods:

- a. When the agent (the person or thing that does the action) is identified in the sentence, put this agent at the beginning of the sentence, as the subject. The subject must always be the noun that does the action of the sentence, as shown in the diagram below.

Example:

WRITE: A switching relay connects the circuits.

(Active)

NOT: The circuits are connected by a switching relay. (Passive)

- b. Change an infinitive verb to an active verb.

Example:

WRITE: The computer calculates the altitude from these values. (Active)

NOT: These values are used by the computer to calculate the altitude. (Passive)

- c. In procedures, change the verb to the imperative ("commanding") form.

Examples:

WRITE: Continue the test. (Active)

NOT: The test can be continued by the operator. (Passive)

WRITE: Remove oil and grease with a degreasing agent. (Active)

NOT: Oil and grease are to be removed with a degreasing agent. (Passive)

d. In descriptive text, when the agent (the person or thing that does the action) is not identified in the sentence, you can use the pronoun subjects “you” or “we” in the active sentence if the agent is the reader (“you”) or the manufacturer (“we”).

Example:

WRITE: On the ground, you can open the valve with the override handle. (Active)

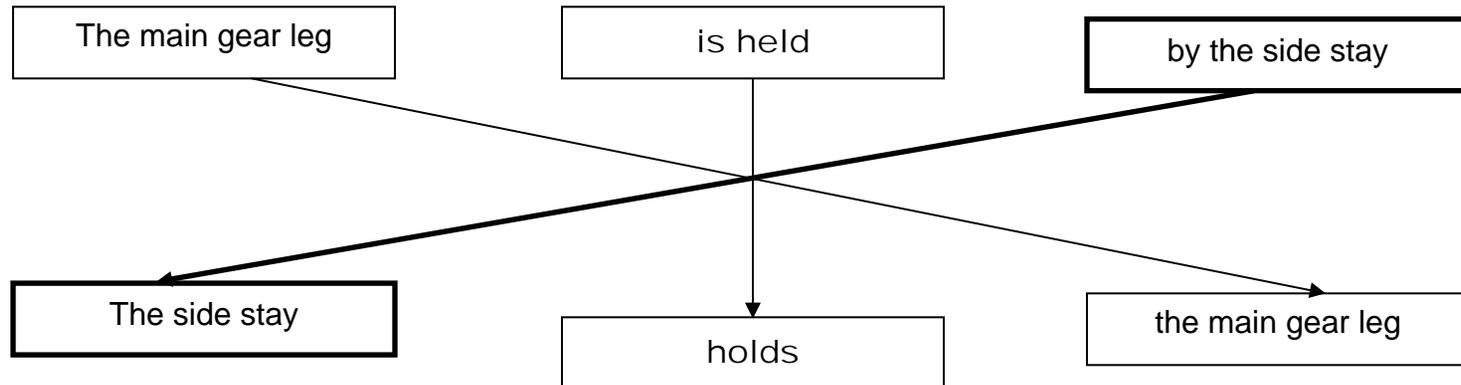
NOT: On the ground, the valve can be opened with the override handle. (Passive)

Example:

WRITE: We do not use the bypass filter now because... (Active)

NOT: The bypass filter is not used in the system because... (Passive)

NOTE: In this case “we” stands for the manufacturer. Note also that the use of personal pronouns can help to make a text more interesting to read.



Wrap-up points on opportunities for human translators:

1. Authoring an original document.
 - a. Keep in mind the approach of “correction as translation.”
 - b. Weigh in on which CE to use.
2. Translation, with and without checkers.
3. Editing and proofing, after batch MT or “HT.”
4. Review glossaries, synonyms at the “analyzing to control terminology” stage.
5. Be involved in the “expanding/customizing” phase, especially if it effects writing, not just vocabulary.
6. You are the best consultant on the TEP process, and it needs to be reviewed before applying any technology decision is made by IT.
 - a. time
 - b. rates
 - c. E & O
7. Be the expert on products, not just CEs, before a purchase decision is made.