

Investment in Quality Pays Huge Dividends



Kent Taylor, acrolinx North America, Inc.



Kent Taylor
Acrolinx North America, Inc.
kent@acrolinx.com

Kent Taylor is a recovering TechPubs veteran, early adopter of SGML, XML, CMS, Single-sourcing, and automated multi-language publishing processes, and Quality zealot. He is currently General Manager of acrolinx North America, Inc.

WHAT'S THE PROBLEM?

Is your company going Global? Already Global? Are the costs of translating required documents and web site content going ballistic? You're not alone—the accelerating pace of globalization and the high (and often unbudgeted) cost of translation impacts more companies and their information development operations every day. And the management of translation costs is becoming an increasingly hot topic of discussion in executive circles around the world.

As the focus on translation costs grows, pressure to reduce costs will quickly shift from localization departments to tech pubs, marketing, methods and procedures, engineering, tech support, and web site information developers. All of the players will be quick to point out that they have modernized their operations, deployed the latest tools, and optimized their processes. They are more productive and efficient, producing greater quantities of deliverable content with fewer resources than ever before. Individually the improvements are generally impressive, but collectively it is unlikely that these independent efficiencies help to control translation costs at the enterprise level.

Each of these departmental optimizations uses different tools, processes, and standards, generally requiring different translation tools, processes, and vendors. It's rare to find a sizable organization that has had the foresight to address these separate information supply chains from a corporate perspective (I haven't found one yet!).

It's a daunting task to even think about pulling all of these diverse operations into a common operating environment.

IS THERE ANY HOPE?

So, what can a company do to address this problem? You might start by looking for what all content supply chains have in common—which at first glance appears to be nothing. But there is one thing that all supply chains have in common and it's so obvious, it's easy to overlook. All of the content development processes include a handoff of source content to a translation/localization provider.

The next logical step is to ask your translation providers what makes translation more expensive than it has to be. The answer will always be the same: the quality of the source language content.

Regardless of how the content was created, it is the content itself that the translators have to deal with.

“The concept is simple: high quality documents that are consistent (in terms of style and terminology), correct (spelling and grammar), and concise routinely cost less than ‘average’ quality documents that are adequate for a native English-speaking audience.”

FOCUS ON QUALITY

The concept is simple: high quality documents that are consistent (in terms of style and terminology), correct (spelling and grammar), and concise routinely cost less than ‘average’ quality documents that are adequate for a native English-speaking audience.

The cost and time savings can be substantial, frequently amounting to a 10-30 percent reduction in translation cost and time for each independent supply chain. It can be even more if you can drive consistent style and terminology across multiple content

supply chains and consolidate your translation processes.

I know this kind of savings may be difficult to believe, so let's break it down into its constituent parts and maybe it will become more believable.

CONCISE

Let's address concise documents first. In a comprehensive independent study of a 1.1 million word software manual set, Mike Dillinger of Dillinger Consulting found that

- ◆ 15 percent of the words were simply not needed. Consider the following:
 - the installation process is very easy (55 percent reduction)
 - connect your PC to the printer with a USB cable (29 percent reduction)
 - There were frequent occurrences of sentences with unnecessary modifiers (words such as very, extremely, exceptionally, approximately, and so on)
- ◆ 50 percent of sentences were at a college reading level or higher (only 20 percent of Americans read at a 10th grade level or higher and non-native speakers typically have a lower average reading level in other than their native language)

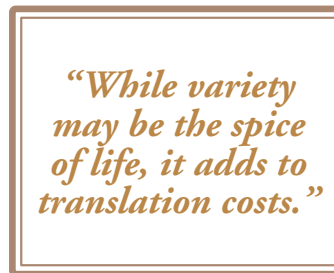
If we were to simplify these complex sentences (45,000±) and eliminate only two words per sentence in the process, we would have reduced the word count by an additional 90,000 words, and the simpler sentences will be easier and less costly to translate.

Looking only at word reduction, that's a total of 255,000 or 23 percent fewer words to send to translation. Multiply that by the number of target languages, and your average rate per word...you get the idea.

Is it worthwhile to focus on being concise? In this example, if you were translating into five or six different languages, you could pay to clean up the English source and translate to an additional language within your original translation budget. And you would still have enough left over for a big bonus for the clever

person who suggested this approach.

Or perhaps you think that this example is atypical, and it couldn't possibly happen in your documents? The 15 percent unnecessary word count in this document set might seem high, but consider this—it only takes 5 out of every 100 sentences with an unnecessary 3-word phrase, or 15 sentences with one unneeded modifier, or some combination of these two ingredients to add up to 15 percent. As a professional writer, you have developed a much larger vocabulary than the majority of your readers (and translators), and you've been trained to use it. However, if your content is going to translation, or to non-native speakers, or even to average English-speaking readers, less really is more.



CONSISTENT

Now let's look at consistency. Obviously it's important to be consistent in the use of trademarks, corporate identity, and branding terminology, but what about 'ordinary' words. By consistently using the same word or phrase to convey a concept or action, you reduce the number of words to be translated and the vocabulary that a non-native speaker (or poor reader) needs to read your documents. A simple example: verify, check, confirm, make sure, be sure, ensure, insure, and assure can all be used more or less interchangeably in most contexts. While variety may be the spice of life, it adds to translation costs. Pick one word and use it consistently. Managing terminology at the word- and phrase-level can be worth another 5 to 10 percent reduction in translation costs, and it can improve readability by a similar amount.

Figure 1 illustrates the impact of terminology management on Translation Memory performance, before and after terminology checking and cleanup. You can

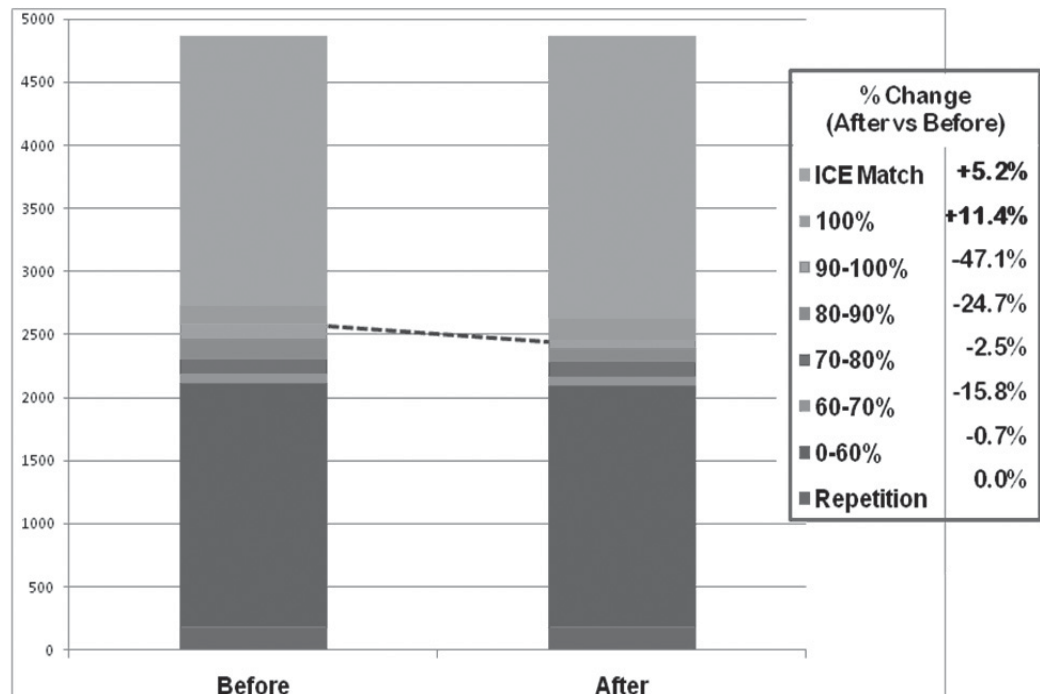


FIGURE 1: TRANSLATION MEMORY PERFORMANCE IMPROVEMENTS

expect similar or better improvements in manual translation environments as well.

CORRECT

We don't have hard metrics on the impact of spelling and grammar issues on translation cost, but we do know that there are a number of grammar issues that cause problems like long sentences, noun strings, prepositional phrases, separated verb parts, split infinitives, and the like. Also, spelling errors that might be incorrectly 'fixed' by a translator can wreak havoc.

Still don't believe that a focus on the quality of your content can reduce costs? Ponder the charts in Figure 2. The chart on the left is from an independent study by Trisoft and Comtech Services and illustrates the impact of various technologies on total translation cost. Note that one of the biggest percentage reductions is on "Pre-Translation cleanup"—30 percent. Note also that this is generally the last step in the search for cost reduction.

The chart on the right is our view of the world at acrolinx. All we've done is change the order of implementation. Start by improving the quality and reducing the quantity of your

content, and feed this clean content into Translation Memory. Neither of these steps requires any change to your existing authoring/editing environment or processes. Managing the quality of your source text can reduce the cost of translations by an average of 25 percent, and the use of a clean Translation Memory can save another 25 percent on subsequent translations. When you decide to move to XML and Content Management, the conversion will go more smoothly with your clean, high-quality content.

OPERATIONAL BENEFITS OF SOURCE CONTENT QUALITY MANAGEMENT

In addition to translation cost reduction benefits, a focus on the quality of your source content pays operational dividends. As every Six-Sigma Black Belt knows, correcting errors early in the process costs a lot less than finding and correcting errors further downstream.

Figure 3 (by J Schütz & R Nübel, IAI) illustrates the relative cost of fixing errors at various points in a typical content supply chain. If errors that are caught and corrected by the author cost \$1, the same error would cost \$2 if it was caught in editing, \$5 if caught in acceptance/release approval, \$10 if caught

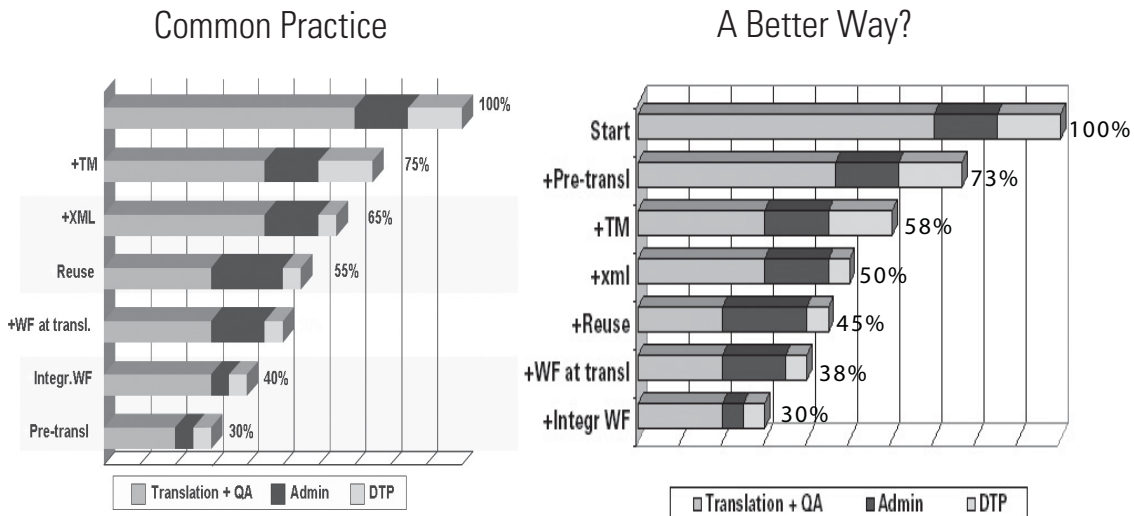


FIGURE 2: TOTAL TRANSLATION COST - TWO DIFFERENT VIEWS

in translation, and \$20 (per language) if it gets caught in the field. At the other end of the spectrum, if you can get your SMEs to provide clean, consistent source material, you wouldn't need to spend \$1 fixing their ten cent errors.

HOW CAN YOU MANAGE CONTENT QUALITY?

It's one thing to identify the problem/opportunity, but when it comes to the solution, the devil is in the details. The traditional methods of written style and terminology guides, initial and recurrent training and heavy editing, can only go so far. Without real-time feedback to the authors, infrequently applied style rules

and/or approved terms are quickly forgotten. Deadline pressures rarely allow time for frequent reference to style guides or glossaries—even if they are available online.

But hope is on the horizon. New software tools are emerging that can guide authors to create higher quality, more easily translatable content. Spelling and grammar checkers are a partial answer but have their shortcomings (which are well known to you).

Controlled Language checkers are another option, but most are based on the highly restrictive rules and 971-word vocabulary of the ASD standard for Simplified Technical English. The standard is a joint development of

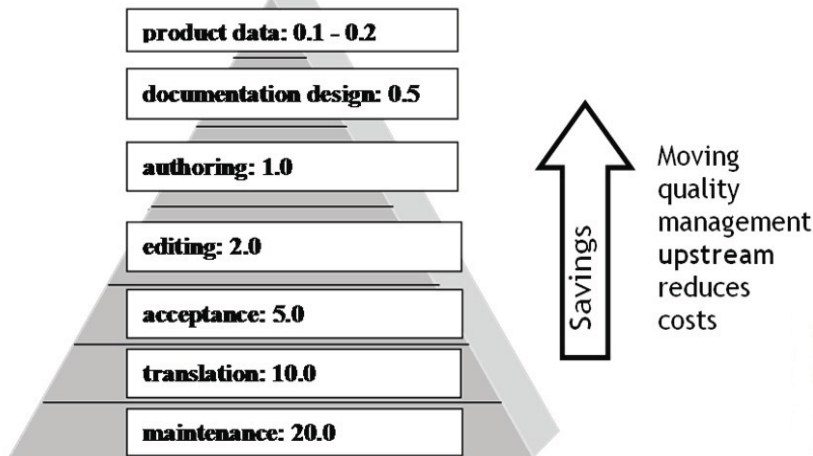


FIGURE 3: RELATIVE COST OF FIXING CONTENT ERRORS

airframe manufacturers and airline companies, designed to eliminate the need for translations. The manufacturers teach their writers to write according to the rules, and the airlines teach their employees the limited sentence construction rules and vocabulary necessary to read the STE content.

The last option is a content quality management tool. The tool of choice of a large and growing number of Global 1000 enterprises is *acrocheck*[™] from *acrolinx*. With a sophisticated Natural Language Processing engine under the hood, and a spellcheck-like user interface that plugs in to Word, FrameMaker, Arbortext Editor, XMetaL, and AuthorIT, *acrocheck* does essentially everything a human copy editor does, except argue with writers and SMEs. It quickly analyzes text for

- ◆ spelling errors
- ◆ grammar issues
- ◆ translatability issues
- ◆ conformance to style rules
- ◆ consistent use of approved corporate terminology
- ◆ opportunities for reuse at phrase- and sentence-levels

Writers and editors get real-time feedback (quality assurance) via different colored highlights that mark suspect text. Clicking on a highlight pops up a dialog box with a brief description of the issue and a suggestion for clearing the issue, with detailed help another click away if needed. This real-time checking capability enables writers to correct errors early in the supply chain. The real-time feedback also performs an on-the-job training function.

For quality control purposes, project managers and quality managers get on-demand reports and metrics at various gates or handoff points in the process to assess quality and to help identify workflow issues. These reports take the form of detailed individual document quality reports, or aggregated reports displaying project-, department-, or enterprise-level quality status.

We find this quality assurance/quality control approach to be quite effective, especially in distributed writing environments. Organizations that deploy *acrocheck* routinely realize the kinds of savings presented in this paper. For more details, go to www.acrolinx.com.

WHAT'S NEXT?

The linguistic engineering team at *acrolinx* is pioneering a new area of sentence-level consistency that promises to have an enormous impact on translation costs. They analyzed a three million segment translation memory with a microclustering utility and found 250,000 'clusters' of linguistically equivalent sentences. These clusters ranged from two sentences to over 100, with an average frequency of 2.5 sentences per cluster. By choosing a single sentence from each cluster to be a 'standard' approved sentence and storing it in a reuse repository, the *acrocheck* engine will be able to identify and flag variant sentences and suggest that the standard sentence be used instead.

If these results are typical, another 10 percent improvement in Translation Memory performance can be expected by the elimination of variants of commonly used sentences. For example, consider the following sentences:

- ◆ Fill the fuel tank at the end of each day of operation to avoid moisture condensation.
- ◆ Refill fuel tank at the end of each day of operation to avoid moisture condensation.
- ◆ To avoid condensation in the fuel tank, fill at the end of each day's operation
- ◆ Fill fuel tank at the end of each day's operation to prevent condensation in tank.
- ◆ Fill the fuel tank at the end of each day's operation to prevent condensation.

would all be highlighted as candidates for replacement by the 'standard' sentence "To avoid condensation of moisture in the fuel tank, fill it at the end of each day of operation." We believe that this minimally invasive support for the writer will bring huge benefits in promoting reuse at very granular levels. 