Context in Translation: definition, access, and teamwork

Appendix 1

More Examples of the Need for Context in Practical Translation

In the body of the article, section B provided examples of each of five aspects of context (co-text, chron-text, rel-text, bi-text, and non-text). Length limitations required the following additional examples to be moved to this appendix. Note that some examples involve more than one of the five aspects.

1. Co-text

Another example of lack of co-text is found in the messages from the Rainbow Portal (http://www.rainbowportal.net/) software system (2005 version) as they are presented to a translator, for the preposition “in.” How does one translate a preposition without knowing what verb it collocates with and what object it has? For example, "in" can correspond in French to "*dans*," "*en*, "*à*," "*depuis*," or even "*avec*", depending on the co-text. Another item lacking co-text is "Lastname." Is this the previous name or the family name?

Even more disturbing is the fact that in this case, even the full XML document available to the translator does not resolve the problem. Not only is the XML not very readable to a translator in its raw form, but the needed context is simply not included and therefore cannot be displayed by any tool. The messages are basically a two-column glossary of an item's name and its value, without any indication of the context of the item.

Figure 1 shows the XML document related to the Rainbow Portal example. It is taken, by permission, from an article on localization quality management by Keiran Dunne (2006).

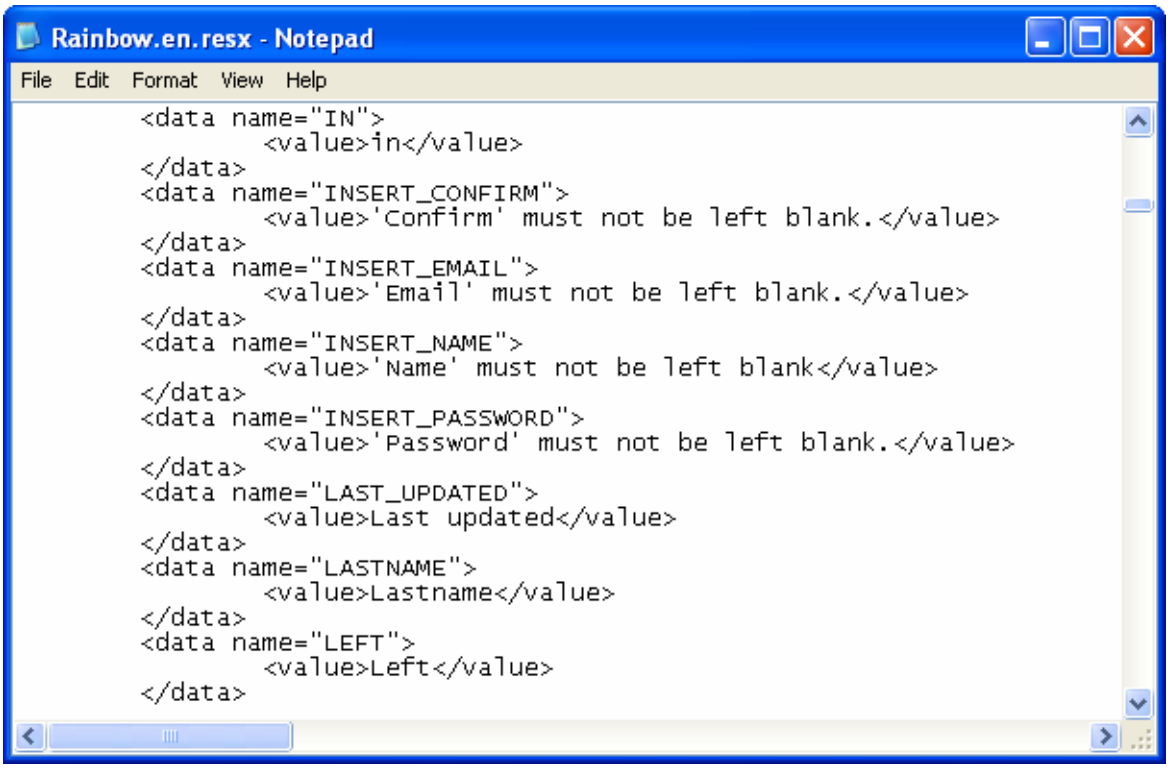

Figure 1 Information Provided to a Translator regarding Software Localization

Clearly, the translator needs access to the co-text in which these strings will occur and the ability to see how translations of the strings will fit together in the localized version of the system. This is possible with some of the most recent software localization tools, but the cooperation of all members of the multilingual communication team is required to give the translator what is needed to see messages in context.

Following are two strings that were presented to a professional translator in an alphabetized list without co-text. Note the ambiguities:

“alert rule settings” (does "rule" group with alert or settings?)

“event console purge policy” (what is being purged?)

A current trend is the use of content management systems (CMSs) – systems that store chunks of text and recombine them in various ways. This practice can cause problems in source text coherence as well as lack of access to context for translators. For content management to work well, the entire approach to writing documentation must be re-thought. The emergence of structured authoring techniques such as DITA (Oasis 2007) is promising in that it promotes authoring that reduces dependence of co-text beyond the immediate chunk. However, it is

imperative to remember that using a markup system such as DITA does not automatically eliminate the need for context and does not necessarily make each chunk of text context-independent.

2. Chron-text

Texts often change over time. Thus, a diachronic view of a source text is needed. However, translation is done synchronically. When a text changes after translation has begun, extra expense is involved, especially when multiple target languages are involved. In one extreme real-life example, a one-word change to the name of a product whose voluminous documentation had already been translated into many languages cost over half a million US dollars for revisions.

Another aspect of chron-text is the shortening of written forms over time. A real text contained the abbreviation "DC" without explanation in the co-text. Human intelligence suggested that the abbreviation did not stand for "direct current," which would be statistically very common. An inquiry revealed that the abbreviation stood for "downloadable characters". At least the abbreviation triggered the need for an inquiry by the translator. A more mechanical approach not involving full understanding of the source text may have used the most common expansion of “DC” without inquiry.

3. Rel-text

In order to make translations maximally understandable, translators must often seek out rel-text to discover what terms have been used to designate them in documents written in the target language or previously translated into the target language. However, it is not always trivial to obtain the needed rel-text. Sometimes real-world knowledge (non-text) is needed in order to identify relevant text or to discard irrelevant text.

Consider the task of translating the following line from a comic strip:

"So… you drank the Kool-Aid" (spoken by a comic strip character).

How should this be translated? Kool-Aid is a powder that is mixed with water to make a flavored drink popular with children. However, neither leaving Kool-Aid in English nor substituting for Kool-Aid a children's drink typical in the target culture will help the target-language reader to interpret the comic strip.

There is only one scene in the comic strip, and it shows two characters interacting, one in a room and the other at the door. The problem is that there is nothing in the drawing of the comic strip to indicate that anyone is drinking anything. Is there something else going on here? One might expect the co-text to

provide an indication of whether to translate the sentence literally; however, the only co-text in the comic strip is this: "Wipe your feet before coming in and don't shed on the sofa". The co-text does not appear to be coherent with the sentence in question. Suppose we do not have access to the author of the comic strip to ask what it means and to ask for previous incremental versions of the text (chron-text) that might shed light on the meaning. We could consult other instances of the comic strip (rel-text) and perhaps infer that the character asking that feet be wiped and hair not be left around is not usually so respectful. A Google search for "Kool-Aid" results in over 250 thousand possibly related texts. What turns out to be relevant is a 1978 tragedy in Guyana where followers of an American, Jim Jones, willingly drank Flavor-Aid (often mistakenly reported as Kool-Aid) laced with cyanide and thus died in a mass suicide. As a consequence, "drink the Kool-Aid" became an expression meaning that one submits unnecessarily to someone else's harmful ideas through blind faith. A piece of rel-text revealing this real world knowledge would allow one to understand the reply in the comic strip ("So… you drank the Kool-Aid") as a rebuke to his colleague for his uncharacteristic submission to the desires of the head of household. Obviously, a direct translation of the remark about Kool-Aid would not be appropriate. Some colorful expression that derides unnecessary submission of authority would need to be found, depending on the audience and purpose of the translated version of the comic strip.

The comic-strip example shows how important rel-text can be and how difficult it can be to recognize the need for additional research rather than accepting a more literal interpretation of an aspect of the source text.

Examples of the need for extensive rel-text need not be as exotic as the comic-strip example. The point is that the translator needs to either intuit or be explicitly told that there is additional, relevant information not found in the source text.

In information technology texts, the following terms have particular meanings that are not easily deduced without background knowledge: "breadcrumbing" (nothing to do with food but based on acquaintance with children's fairy tales), "big-endian" and "little-endian" (not a misspelling of big and little Indians but part of the Unicode specification), and "wizards"(nothing to do with magic but rather a user help in software). Access to rel-text is necessary to properly translate such terms. The rel-text can be highly processed, such as a monolingual terminology database with domain information about each concept, or can be raw text.

Terminology work can be thought of as a process of making rel-text and some aspects of non-text into structured text that is easily accessible.

Clearly, when rel-text is needed but not supplied and the need for rel-text is not detected because of insufficient non-text, the result can be serious.

The availability of re-text is only a partial solution. Intelligent use of rel-text is an aspect of translation competence that develops over time.

4. Bi-text

Large organizations often spend considerable time and money cleaning up translation memories so that they are consistent, both internally and with organization-level terminology standards. Even if care is taken to place only reviewed translation material into a translation memory database, changes in terminology over time may require retroactive modifications to older translation units that are to be used as a basis for new translations. Those that do not maintain their translation memories pay a price when translators use inaccurate translation units and produce translations that need to be corrected. Either way, bi-text and its derivatives are important to the translation process.

Sometimes rel-text (in the language of the source text) and bi-text resources are both missing. Another real-life example of the need for rel-text and bi-text involves an international court case. A US law firm involved in the case found that a large number of court documents in Italian needed to be translated. Some of the documents were entrusted to one translation company and other documents to a second translation company. Both translation companies used free-lance translators, and some free-lance translators were asked to do work for both translation companies. Thus, those translators were able to see differences between the rel-text supplied by the two translation companies. One of the translation companies split up a set of documents among three translators. These translators were given a style guide (a very important resource, of course) containing very specific instructions, such as a rule not to use the word "since" in the target text when it means "because." Then the work of these three translators was given to three separate reviewers, each reviewer examining the work of one translator. Unfortunately, there was no coordination among the reviewers. The reviewers were given no instructions at all, not even the style guide. And no one in the process, neither the translators nor the reviewers, had access to a project glossary. The result, predictably, was a divergent set of documents. It was very expensive to fix the problems caused by lack of rel-text and obtain a consistent set of documents.

The court-case story continues. The other translation company maintained a bilingual terminology database of legal terms (an important type of bi-text) and supplied it to the translators working on their documents. Some Italian legal terms with no exact equivalent in the American legal system were even specified to be left in Italian (with an explanation of the term). This allowed the translators to produce a set of documents with consistent terminology. However, this translation company had no style guide. Thus, there were other types of inconsistencies in the translations. Clearly, the lack of sufficient access to resources (glossaries and style guides) in this case resulted in serious problems.

This painful example is not unusual. The lack of access to needed rel-text and bi-text resources is all too common.

5. Non-text

Non-text, as previously mentioned, includes aspects of the saying that cannot be fully captured in the said. Consider the difficulty of treating idioms that can also be interpreted literally, depending on the context. One example is "bull's eye" (the center of a target). Suppose this expression is found in a passage about hunting bull deer (that is, male deer). Imagine the emotional impact on a hunter of looking at a bull's eye through a telescope sight on a hunting rifle. A hunter who put down his rifle might describe this event to a close friend who, he believes, will not ridicule him for failing to kill the deer when he had an opportunity. A machine might have trouble determining whether this text is referring to the visual organ of a deer or to a paper target nailed to a tree. This is because machines do not currently have access to the saying. It is not clear that they ever will.

Documented errors in movie subtitles illustrate how easy it is to make a mistake if the translator does not have a firm grasp on the source-language culture. A horror film shown in Kuala Lumpur included the subtitle meaning "cereal killer" for the spoken English phrase "serial killer", which sounds the same (http://thestar.com.my/news/story.asp?file=/2007/7/21/nation/18368076&sec=nation).

In a British television show, dialogue meaning "she died in a freak rugby accident" was given the subtitle "she died in a rugby match for people with deformities" (http://entertainment.timesonline.co.uk/tol/arts_and_entertainment/film/article1533920.ece).

The translators who made these mistakes did not give sufficient consideration to non-text. By reconstructing the saying in his or her mind, the translator would probably have been alerted intuitively to the need for more research before submitting these subtitles.

References

Dunne, K. (2006). "Putting the cart behind the horse: Rethinking localization quality management" in Keiran Dunne (ed.) *Perspectives on Localization*, (pp. 95-122). Amsterdam/Philadelphia: John Benjamins.

Oasis. (2007). Darwin Information Typing Architecture (DITA) Version 1.1 (http://docs.oasis-open.org/dita/v1.1/overview/overview.html)