

Embedding Technology at the Front End of a Human Translation Workflow: An NVTC Vision

Carol Van Ess-Dykema*, Helen G. Gigley‡, Stephen Lewis*,
Emily Vancho Bannister*

National Virtual Translation Center
Washington, DC 20535 USA
cvanessdykema@nvtc.gov, hgigley@gmail.com,
stephen.p.lewis@ugov.gov,
emily.c.vancho@ugov.gov

Abstract

This paper describes the strategic vision for a new translation management workflow for the US Government's National Virtual Translation Center (NVTC). The paper also describes past, current, and planned experiments validating the vision, along with experiment results to-date. The most salient features of the new workflow include the embedding of translation technology at the front end of the workflow (e.g., translation memory technology, specialized lexicons, and machine translation), technology-generated "seed translation", a new human work role called "paralinguist" to assess the "seed translation" and assign an appropriate translator/post-editor, and new human translation strategies including federated search of online dictionaries and collaborative translation.

1 Introduction

Established by the United States Congress in response to 9/11, the National Virtual Translation Center (NVTC) has as its mission to augment the foreign language translation capabilities of the Intelligence Community (IC) and Department of Defense (DoD). Its translators work both on-site and virtually, and soon NVTC will provide 24/7 translation service.

* On rotational assignment from the Department of Defense

‡ On rotational assignment from OCS

This paper describes NVTC's strategy to respond to its customers' translation requests with greater speed and accuracy by:

- Validating a translation management workflow that embeds emerging technology at the front end, generating an initial "seed translation" of documents for human post-editing
- Conducting experiments that measure the utility of emerging technologies on NVTC translation languages, genres, domains and media
- Establishing business processes that will make NVTC the most cost-effective IC element for human foreign language translation

2 Traditional Human Translation Workflow

Figure 1a portrays the traditional human translation workflow, both in government and in industry.

The first significant aspect of this workflow is that each human translator traditionally uses his or her self-developed/maintained wordlists which he/she shares with few other translators. Although major government agencies are able to provide corporate lexical resources to their translators, development and use of corporate linguistic knowledge is often minimal in smaller agencies.

Second, the traditional workflow is sequential. Translators begin translating from the original source document. The translators complete the translations before the quality control process begins. The quality control professional potentially

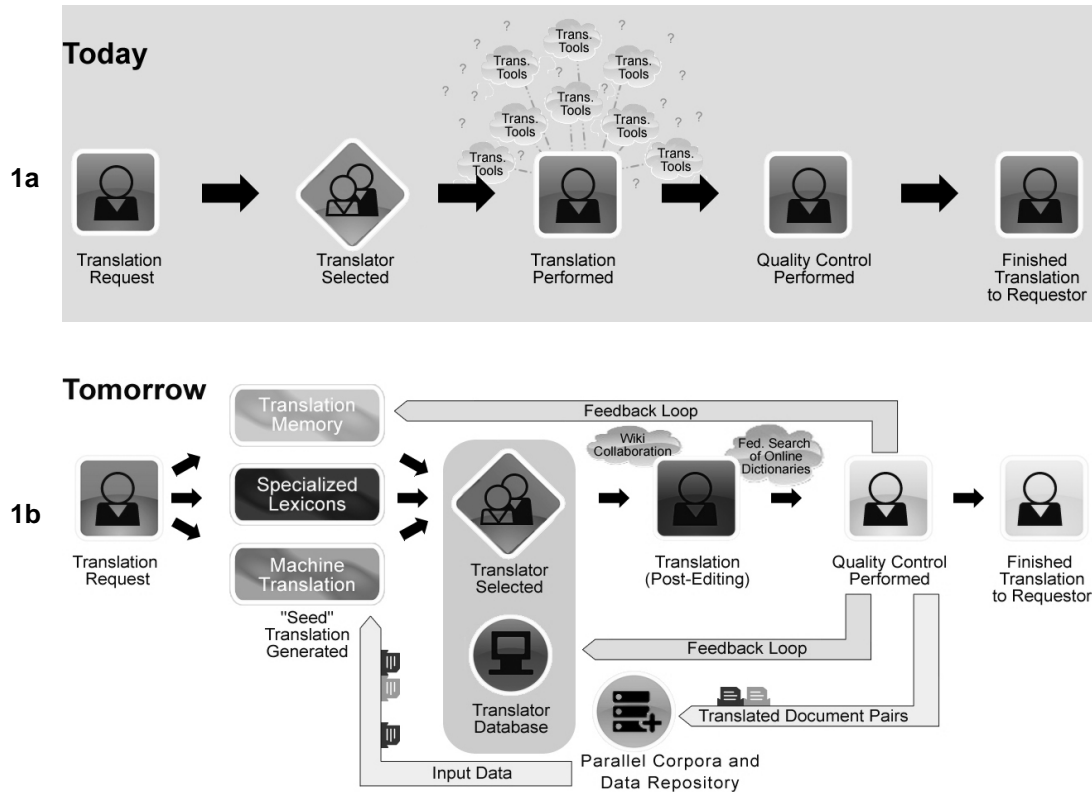


Figure 1: NVTC human translation workflow vision

must repeatedly verify the necessity of a word change throughout the target document.

3 Developing a New Translation Management Workflow¹

NVTC has proposed a new workflow exhibited in Figure 1b.

The most salient features of the new workflow are the following:

- The insertion of the technology at the front end of the workflow
- Technology-generated “seed translation”
- A new human work role called “paralinguist” to assess the “seed translation” and to assign an appropriate translator/post-editor

- New human translation strategies, including federated search of online dictionaries and collaborative translation
- Feedback loops for continuous improvement of the human translation workflow

Specifically, the “seed translation” is automatically generated from one or more of the following technologies, that we envision being accessible through a service oriented architecture:

- Translation memory (TM) technology
- Specialized lexicons
- Machine translation (MT)

The roles of the “paralinguist” are:

- To estimate the time and cost for the post-editor to produce the final translation from the “seed translation”
- To select the best translator/post-editor based on database records of translator-specific speed and accuracy

The new online strategies for the translator/post-editor of the “seed translation” include the following:

¹ During November 2007, NVTC Directors traveled to Ottawa to meet with their Canadian Government translation counterparts, the Canadian Federal Translation Bureau (CFTB). The CFTB shared its workflow with NVTC, noting 1) the embedding of translation technologies into the front end of the workflow and 2) a resultant, significant return on investment for speed.

- Federated search of online lexicons/dictionaries
- Wiki translator collaboration tools and quality control

The three feedback loops include:

- Translated document pairs (sentence aligned) loop back to the parallel corpora and data repository for input to machine translation development.
- Aligned translation document pairs loop back to translation memory banks.
- Translator quality metrics (speed and accuracy) loop back to the translator data base.

4 Validating the NVTC-Proposed Translation Management Workflow

The NVTC translation management workflow in use today includes “hooks” for translation technologies. NVTC is assessing the output from different types of technologies to determine which to incorporate into its translation workflow.

4.1 Synergistic Workflow Modules

Zetzsche (2007) reports that modules within translation environment tools are an expected and required component, with translation memory (TM) and terminology requirements often having to fit into a larger framework of process management.

He also notes that it is in the interest of machine translation (MT) tools to have a strong translation memory and terminology management component, and that it is equally in the interest of translation environment tools to offer an easy and seamless interface and integration to machine translation.

4.2 Assessing a TM tool with workflow management capability

NVTC is currently assessing a combined translation memory and workflow management technology system for its potential improvement to the NVTC workflow.

The assessment of the system is based on written material that describes a translation memory system whose workflow includes bilingual translator dictation of the target language and subsequent monolingual post-editing of the target text.

NVTC’s initial assessment of this workflow is that it is unsatisfactory for NVTC because NVTC

translators are unaccustomed to translating orally from the source language into the target language and because NVTC quality control editors tend to be other bilingual translators, not monolingual target language speakers.

From a business process perspective, this system is additionally inadequate: NVTC is unable to hire the number of monolingual post-editors for the large number of languages it translates.

NVTC seeks to identify and assess additional translation management workflow systems that include translation memory capability (see section 6).

5 Transcription and Translation of Multimedia

As already noted, NVTC translates a variety of languages. It also translates a variety of genres, domains and media. It has recently assessed an automatic transcription tool for its utility in translating multimedia captured video and audio on-line broadcast feeds.

5.1 Results from an assessment of the utility of machine-generated transcripts for the task of translating TV news broadcasts

During 2007 NVTC worked with US Government personnel at the National Center for Language and Culture Research and its research scientists at the University of Maryland Center for Advanced Study of Language (CASL) to measure improvements in translation speed and accuracy that result from utilizing an automated aid, namely transcription, versus creating an un-aided translation.

Potential benefits differ for native and non-native speakers of Arabic and Mandarin, as determined in CASL’s assessment of the transcription tool. The study additionally systematically varied three key tool dimensions in order to ascertain their importance independently of the tool itself:

- Availability of a machine-generated transcript
- Availability of navigation features embedded in the transcript
- Transcript quality

A total of 54 Arabic translators and 54 Mandarin translators (split evenly between native and non-native speakers in each language) participated. (Powell and Blodgett, 2007)

The findings indicate the following:

- Native Arabic translators gained little in translating using the transcripts but maintained their productivity.
- A percentage of the translators felt the transcripts were of little help due to errors in them and preferred to translate directly from the broadcast byte.
- Some indicated they would like to be able to edit the transcriptions.
- Non-native Arabic translators made fewer errors when having the transcripts available.
- Having both the linked broadcast and transcript with highlights provided the most help to the non-native translators.

The research used a compound-computed score to measure time to translate and quality of the translation.

5.2 Assessing machine-generated transcription and machine translation for the task of translating TV news broadcasts

The 2007 CASL findings provide documentation of potential roles that transcripts can play during translation and suggest which translators they may serve best. As a follow-on, the NVTC realized that this finding may be useful in identifying a set of translator aids that should be provided to all translators, especially those who are Independent Contract Linguists. These are freelance linguists who contract directly with the NVTC, rather than through a vendor. They reside in a virtual translation space and translate wherever they are located.

NVTC would like to assess the needs of this set of translators, but in the real environment, via web access. The experimental conditions will include all those of the previous experiment except the unlinked transcription and broadcast byte condition. Instead, a condition using the machine produced translation linked with the transcription and broadcast byte is to be used.

Of interest in this condition, is the effectiveness, or lack thereof, of the provided translation. It is suggested that one complete set of the original stimuli that was used in the linked broadcast byte and transcription evaluation be used for cross-condition analysis but with translators not in the first study. Metrics on time to translate and quality of the translation, and feedback from the remote

translators on how they felt the tool aided them and in what ways, will also be studied.

The results from the study will inform NVTC about the virtual environment its translators use and will suggest the scope of translator aids they may wish to supply to all translators. Investment in resources may be cost effective in the long run. Both time to translate and amount of time for quality control impact the cost to produce a translation. As before, NVTC is seeking to understand if time to translate can be reduced without impacting quality of translation. The hope is that the quality can also improve. This will reduce quality control time to produce the final document.

6 TM Needs for Multiple Genres

Translation memory systems use collections of paired sentences, a source segment and its translation, to help translators more readily use previously translated text. The typical TM system automatically queries a database of previously translated sentences while a translator moves through a translation task. As matches are found, the translator can simply re-use the translation and need not re-translate any segments already translated.

Since TM systems are designed to facilitate translation re-use, they are most effective when used with translation tasks that involve repetition, such as updates to product manuals and documentation. In translation of documents involving minimal repetition from previous documents, TM systems have had much more limited success.

6.1 NVTC experiments to assess TM with multiple genres

TM system developers typically use counts of leveraged translation segments to estimate cost savings for their customers. Such measures are dependent upon existing databases. Without an existing translation memory database, there is, to our knowledge, currently no reliable method for estimating cost savings of using TM technology.

NVTC often translates documents that have only minimal repetition from previous translations. Accordingly, NVTC has proposed and has received funding from the Office of the Director of National Intelligence to perform a translation memory experiment to help evaluate the feasibility of using TM technology for the kind of data that NVTC needs to translate.

The NVTC is spearheading a government inter-agency group to learn how other government agencies use TM technology and to ground the NVTC exploration of TM technology in a community of users. Like NVTC, many government entities need to translate a variety of languages and genres.

The first NVTC TM experiment will involve translation of 3 documents in each of 2 languages with two translators for each language and with documents selected from multiple genres. The efficiency of professional translators with the aid of a TM system will be observed. The baseline is provided from documents that have already been translated using NVTC translation and quality control processes. This small-scale experiment is intended to give understanding and direction to guide further studies, including the development of appropriate metrics. We hope that the qualitative human observations on the cross-language, cross-genre utility of TM technology will inform quantitative (statistical) measures.

After gaining an understanding of the benefit of TM software in the NVTC context, NVTC hopes to provide TM technology at the front end of the translation process and also make it possible for translators to use TM software throughout the translation process.

7 Specialized Lexicons

The use of dictionaries has always been the underpinning of the translation process. Domain specific lexicons are essential to maintain consistency and quality of translation in recognized domains. Industries often develop their own preferred lexicons, and many localization vendors provide lexicon creation services. NVTC will partner with other government agencies to select appropriate electronic specialized lexicons for generating “seed translations.”

8 Machine Translation

In this section, we describe an automotive industry use of customized MT and a potential NVTC use of customized MT.

8.1 Use of customized MT by the automotive industry

Rychtyekj (2007) reports that Ford Motor Company has been using MT since 1998. Ford has

translated more than 7 million records describing build instructions for vehicle assembly at its plants in Europe, Mexico, and South America. It uses a controlled language in its source texts; it also translates free-form text comments that are embedded within the assembly instructions.

Rychtyekj (2007) notes that the most difficult issue in developing the Ford translation system was constructing the technical glossaries that describe the manufacturing and engineering terminology which Ford uses - a customized translation system coupled with a set of Ford-specific dictionaries.

Ford also uses a web-based tool that provides its translators the capability to test and update the technical glossaries as needed.

8.2 NVTC Vision for Customized MT

Although NVTC does not necessarily translate the kind of structured documents successfully translated by Ford using MT, NVTC hopes to make good use of MT solutions. In the localization industry, project managers can often read the source document and make reasoned judgments about appropriate translators and similarity to previous documents. At the NVTC, the variety of source languages encountered often makes this difficult. By embedding MT at the front end, “paralinguists” and project managers can achieve some level of understanding in unfamiliar languages, and can make the decisions necessary for directing high quality translation.

As MT technology continues to improve, integrating MT at the front end of our processes will provide our translators a “seed translation” to post-edit.

9 New Translation Strategies

In this section, we describe two new human translation strategies that NVTC is investigating.

9.1 NVTC Federated Search of Online Dictionaries

NVTC is currently assessing a government-produced online dictionary suite for use by its Independent Contract Linguists (ICLs). The dictionary suite is available at all three levels of classification (UNCLASSIFIED, SECRET, and TOP SECRET) and provides a search of multiple

bilingual dictionaries. It is available to government employees and contractors.

Federated search is a strategy for searching multiple heterogeneous databases using a single query. For the future translator, the strategy will query all available dictionaries and return the responses to the translator in a single presentation. With non-federated searching, translators are required to enter the query into each online dictionary individually and wait for each result independently. With a federated search, translators type the search term once and receive definitions from all the dictionaries that contain that term. The simultaneous display of all the search results allows the translator to more easily glean information about the term's connotations and usage.

9.2 NVTC Collaborative Translation

NVTC currently handles large translation projects by providing different portions of the project to different translators. This allows translators to work on different portions simultaneously. However, it doesn't always produce a cohesive, consistent product, as different translators often translate the same phrase in slightly different ways. Collaborative translation aims to improve upon this existing process by encouraging translators to discuss and agree upon lexical usage in their joint work.

9.2.1 Collaborative Translation and Lexical Usage

Lexical usage among translators is variable and fluid and requires constant translator input to maintain consistency. NVTC envisions that its virtual translators will consult one another using a wiki environment or another collaborative platform to allow them to add their own translations of phrases that show up repeatedly in the project. If the translators disagree on how a given phrase should be translated, they can discuss it and come to a consensus.

The types of NVTC translation tasks that will potentially benefit from collaborative translation include books (with many chapters), groups of documents which share a genre (e.g., a group of websites), and groups of documents which share a domain (e.g., documents about weapons of mass destruction).

9.2.2 Collaborative Translation and Quality Control (QC)

In addition to lexical usage, quality control is another place where collaborative platforms can be used to good effect. Currently, NVTC projects are handled by assigning a task to a translator, waiting for the translator to finish the task, and then delivering the final product to the quality control professional.

With translators frequently submitting their work to the collaborative platform (e.g., every night, or even every few hours), the QC professional has much more opportunity to review the work and make necessary changes. When necessary changes are identified early in the translation process, it precludes repeated changes on the part of the QC professional, resulting in the production of a high quality translation in less time.

10 Continuous Improvement

In this section, we discuss two methods for improving the human translation workflow that stem from the quality control step in the workflow.

10.1 Parallel Corpora Alignment (PCA)

The NVTC Parallel Corpora Alignment (PCA) proof-of-concept system is a web application designed to process translation output in a format that will ensure its availability for reuse. The system segments and aligns collections of Arabic-language texts and their English translations at the sentence level. After extensive human editing, the aligned sentence pairs and associated metadata are to be formatted and stored in a repository for possible future use. The only language pair supported by the current system is Arabic-English. The system consists of both automatic alignment and a user intervention environment to ensure ultimate quality.

10.2 Feedback Loops

Workflow feedback loops also provide continuous improvement:

- Translated document pairs loop back to the parallel corpora and data repository for input to machine translation development.
- Aligned translation document pairs loop back to translation memory banks.

- Translator quality metrics (speed and accuracy) loop back to the translator database.

11 Future Directions

This paper presents the NVTC human translation workflow vision and beginning steps in its validation. We seek suggestions from the US Government translator and translation vendor communities that will assist us on our journey.

Acknowledgments

NVTC acknowledges the Technology Development Group effort to develop the parallel corpora alignment proof-of-concept system.

References

- Powell, Allison, and Allison Blodgett. 2007. Using Technology to Aid Translators — Final Results of the Translator's Aide Evaluation. University of Maryland Center for Advanced Study of Language, TTO 301. (Data from Fact Sheet, April 2008.)
- Rychtycky, Nestor. 2007. Machine Translation for Manufacturing: A Case Study At Ford Motor Company. *AI Magazine*, Vol. 28, No. 3 (Fall 2007), pages 31 – 43.
- Zetzsche, Joel. 2007. Translation memory: state of the technology. *Multilingual Computing*, September 2007, pages 34 – 38.