Conceptualising the future of translation with localisation

Minako O'Hagan

Minako O'Hagan, School of Applied Language and Intercultural Studies, Dublin City University

The widespread use of electronic communications modes for international exchanges is making localisation relevant to the emerging need for new language support. Indeed, the recent demand for enabling Web sites in multilingual versions has established localisation as integral to the modern translation business. As a consequence, conventional translation is being increasingly replaced by what the author terms *teletranslation* as the means to enable *translation-mediated communication (TMC)* in electronic modes. This paper explores how elements of localisation practice are providing the basic building blocks of teletranslation. The author therefore maintains that localisation provides insights into the future of translation.

1. Introduction

The birth of localisation coincided with the opening of international markets for the computer industry that began to emerge in the 1980's (Esselink, 2000). Although the single largest component of localisation still is translation, the process of localisation as a whole diverges greatly from conventional translation as the latter is primarily based on print media which is the hallmark of the industrial society. Software localisation involves a number of linguistic and cultural adjustments, concerning not only documentation in print, but software itself. The task therefore is a combination of translation in its conventional sense and software engineering. Many early localisation companies sprung up from the software engineering sector rather than conventional translation operators which found the engineering dimension beyond their scope. This origin has meant that localisation tended to create its own path, more akin to software engineering and separate from conventional translation. However, the 1990's saw a sea change in this notion as the localisation industry firmly established itself to be a significant segment of the modern translation business. The clearest contribution was localisation of the World Wide Web (Web).

In the advent of the Internet and its requirements to enable multilingual environments, new modes of language support emerged. One was localisation of Web sites, enabling the user to view and navigate the content of a given Web site in his or her language. Sophisticated Web sites are now offered in multilingual versions and updated simultaneously across all languages. Enabling and maintaining Web sites in a number of languages was almost immediately established as a localisation task rather than that of conventional translation because this involved software engineering in addition to translation. The localisation industry established best practices and led the translation industry into this new era of multilingual Internet. Web localisation cemented the position of localisation as being indispensable for meeting the emerging need for language support on electronic communications platforms.

Needless to say, however, not all Web sites are localised into all languages and this allowed another popular language support to develop on the Internet in the form of online machine translation (MT). Today many portals and search engines integrate MT as part of their basic utilities to enable the user to obtain the required information in his or her chosen language. The rationale for employing MT was clearly linked to the need for (near) real-time translation, seamless integration of language support into the user's online environment and also cost consideration in line with the 'free information' culture still prevalent on the Internet. To this end, MT services on the Internet are mostly provided free of charge. MT-based language support is well suited to meeting the needs that arise from the Web for low cost real-time information 'jisting'.

As these two examples illustrate, the user environment in which language support is required is changing. Conventional translation, characterised as an asynchronous process catering to print-based text for physical distribution, is becoming less and less compatible with the rapidly expanding ICT-based communications infrastructure. It seems reasonable then to assume that future translation will be increasingly sought in the context of electronic modes of communications based on digital technology. The author calls the emerging language support teletranslation (O'Hagan, 1996; O'Hagan & Ashworth, 2002) to highlight the change from conventional translation. This change will not mean that conventional translation will disappear, but rather, new modes of language support will subsume the function of the former. The term teletranslation is used to mean language support provided to enable interlingual communications in electronic modes whereas the term conventional translation refers to language support in conventional non-electronic modes such as translation of text to be used for offline print media.

2. New Contexts of Translation-mediated Communication (TMC)

2.1 Definition of TMC

In its broadest sense the term *Translation-mediated Communication (TMC)* can encompass any interlingual exchange enabled by the assistance of a translator or an interpreter. For example, a translation of foreign literature, a subtitled foreign film and a meeting assisted by an interpreter can all be regarded as *TMC*. The word *Translation* in *Translation-mediated Communication* includes interpretation as well as translation. Further, in the context of this paper, the main focus of *TMC*, as in O'Hagan & Ashworth (2002), is interlingual exchanges occurring in electronic modes such as CMC (Computer-mediated Communication) rather than the conventional modes based on physical print media or face-to-face interactions. The term *TMC* in this paper therefore takes the narrow meaning of translator- or interpreter-assisted electronic communications. On the basis of Shannon's (1949) communication model, *TMC* can be examined in terms of the *Sender* of the Message in the Source Language (SL) and its *Receiver* in the Target Language (TL) communicating via *Translator* (either translator or interpreter) who converts the Message from SL to TL. In particular, the *TMC* framework focuses on the role of technology by analysing its impact on the *Sender*, the *Receiver* and the *Message* itself in a qualitative man-

ner, thus different from Shannon's model, which is focused on the transmission function of telecommunications in a quantitative sense. The set of terminology with capitalised first letters (*Sender, Receiver, Message* and *Translator*) are used throughout this paper whenever they are referred to in the context of *TMC*. Nida and Taber (1969) highlighted the role played by the translator on the basis of the Shannon's model and this is relevant to illustrate *TMC*.

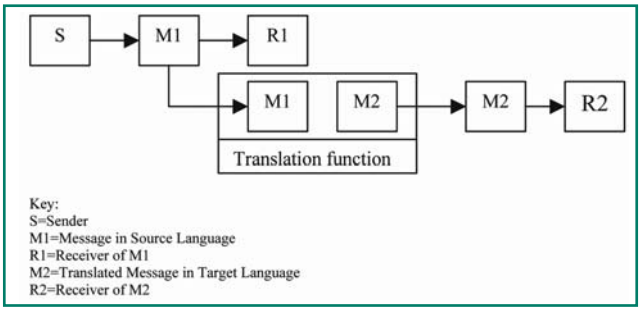

Figure 1. Translation-mediated Communication (modified from Nida & Taber, 1969)

The fundamental function of the translator as the *Receiver* of the *Message* in SL and the *Sender* in TL remains the same between conventional translation and *teletranslation*, but the way in which the *Message* is transmitted, stored and processed show marked differences. The following section takes the case of localised Web sites and examines it in the framework of *TMC*.

2.2 Localised Web sites as TMC

In the late 1990's multilingual Web sites started to emerge in recognition of the nature of this communication platform being immediately global. This created a new domain of language support which requires localisation of a Web site into given locales. In order to highlight the difference from print-based text subject to conventional translation, the following diagram shows the typical lifecycle of Web content also in relation to the *Sender* and the *Receiver* of the *Message*.

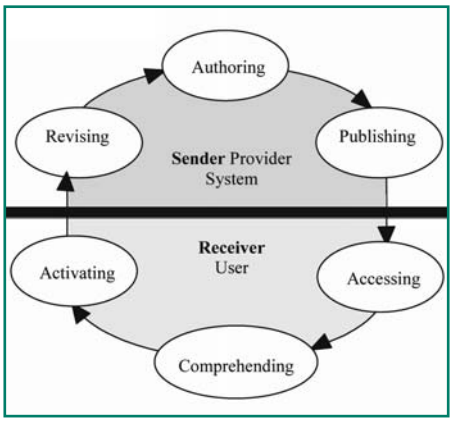

Figure 2: Web content lifecycle (modified from Lockwood, 1998)

The Web allows its users to disseminate and assimilate information in an interactive manner enabled by hypertext. The lifecycle of content for a Web site starts with authoring, which may range from entirely text-based to incorporating audio, moving images and other non-textual elements. Once the site is published by the *Sender*, it goes global unless access is limited such as on an intranet. The *Receiver* accesses the site by some kind of IT device which today includes mobile phones and PDAs (Portable Digital Assistant) in addition to desktop computer terminals. Once the *Receiver* comprehends the content, he or she may act upon the received information. For example, the *Receiver* may decide to place a bid on an item in an auction site. The *Sender* of the Web site is normally able to gather various user information automatically and may use such feedback to update the site. The cycle continues in this manner. The localisation process is typically applied between the authoring and the publishing stages although an application of the internationalisation process will, in effect, advance the overall localisation planning to prior to authoring. Internationalisation is further elaborated in the next section. Because the whole lifecycle is embedded in the digital environment, all aspects of information storing, processing and transmitting of Web content are conducted by computer. Given the frequent updates required of the information, ongoing content management is a significant dimension of maintaining localised Web sites. This is similar to the continuous cycle of software versioning, but is new to conventional translation which generally deals with text not requiring such frequent updates. In the *TMC* framework, a localised Web site can be characterised as follows:

The *Sender* in SL: Originator of the Message.

(1) The *Sender* is increasingly aware that once the site goes online, anybody can access it (unless access is deliberately restricted) and the circulation tends to be wider than that for offline print-based translations.

(2) Web authoring has become a specialised technical task as it involves understanding the characteristics of the Web. For example, readability of the content as well as its aesthetics on screen rather than on paper needs to be taken into consideration. Furthermore, the *Receiver* environments such as operating systems, browsers, character encoding, etc. also become relevant factors to consider.

(3) The application of internationalisation may affect the *Sender* to allow for subsequent translatability of the content.

The *Receiver* in TL: Unspecified recipient of the Web site.

(1) Localised Web sites will be exposed to a wide range of TL readership and accessed via an increasing range of devices, including mobile phones.

(2) Interactivity based on hypertext allows the Receiver to arrive at different parts of the text, not necessarily in the order in which it is originally written.

(3) The *Receiver* will soon discover if the given site is user-friendly in terms of functionality as well as linguistic and cultural dimensions. A localised Web site facilitates the Receiver to understand the content, but depending on the extent of localisation, not all information or functionality may be available in TL. Some Web sites are only partially localised and this may prompt the Receiver to apply ad hoc language support such as online MT.

The *Message*: Web content

(1) In addition to text, Web contents often include non-textual components which may require adapting to the Target culture.

(2) Web content needs to allow for the fact that information is often scanned quickly on screen by the *Receiver*, at least in the first instance to gather relevant information.

(3) The increasing use of mobile devices to access the Web has meant that the content needs to be further tailored to smaller screens of such devices.

(4) Web content is normally updated at much more frequent intervals than that of offline print media.

The *Translator:* Web localiser

(1) The task of converting a Web site from SL to TL involves translation of text in a conventional sense and software engineering.

(2) The *Translator* needs to be aware that the Web content is normally exposed to much wider TL readership than offline text, thus the content has to sound and look natural to a wide range of TL native speakers.

(3) It is essential for the *Translator* to understand the characteristics of the Web as described earlier.

(4) The *Translator* will have to be familiar with content management requirements and be required to use certain translation tools such as Translation Memory to facilitate efficient updating of the content.

A number of differences from conventional translation are obvious. The impact of the new mode of communication on the overall translation process is clear in response to the changes in the characteristics of the *Message*, how the *Sender* has to present the *Message* and how it is consumed by the *Receiver*. The overall process involved in localising a Web site is akin to software localisation in that the task is possible only if supported by computer to fit into the digital lifecycle of Web content. For example, unlike conventional translation, quality control involves computer-based testing in terms of the user environment as in localisation of software. The use of certain computer-based tools is also essential — particularly in view of the ongoing maintenance of the content, which results in frequent changes for updates. In what follows, an emerging example of *teletranslation* is explored.

3. Emerging teletranslation practice: DVD localisation

The localisation process has evolved in response to the need for language support in new technological environments such as computer software and the Internet. With the widespread use of the Internet, localisation of Web sites became the fastest growing area within the translation sector in the late 1990s (Lockwood, 1999). Similarly, the growing use of DVD for audiovisual content may see this medium becoming the next big wave of localisation. DVDs offer enhanced storage and information processing capacities. For example, as compared with CD-ROM which has a storage capacity around 700MB, a single DVD has 4.7 to 17 GB and is able to embed subtitles in up to 32 languages or up to 4 dubbed versions (Karamitroglou, 1999) with both subtitles and dubbed versions on the same disk. Furthermore, DVDs allow interactive features such as scene-based searches, various processing possibilities of selected scenes and user options to show/hide or select subtitles. The capacity of DVDs to support multilingual speech and text will make them an ideal platform for audiovisual content distribution in the global market. The implications of this could be that certain localisation practices such as simultaneous shipping (simship) may become applicable to DVD releases in the same way as for popular computer software and Web sites. This will likely mean that screen translation becomes subject to certain standard localisation processes such as project management, workflow and quality control procedures. Furthermore, the subtitling and dubbing processes themselves may be affected. Given the new interactive features as well as processing and storage capacities pertinent to DVDs, a new mode of screen translation could develop. For example, today's subtitles as a linear text stuck at the bottom of the screen could change to incorporate hypertext and multimodal elements, even combining subtitles and dubbing techniques together so as to optimise the effectiveness of language support.

On the basis of the *TMC* framework, the communication process of a subtitled audiovisual content can be seen as: the audiovisual content (the *Message*) which was originally produced in SL is augmented by the subtitler (the *Translator*) who produces subtitles in TL for the TL audience (the *Receiver*) that enable the equivalent impact intended by the director/producer of the content (the *Sender*) on the SL audience. In this, DVD as a medium is likely to affect the nature of the *Message* as well as the *Sender* and the *Receiver* and also the *Translator*. DVD provides durable digital storage of the audiovisual content and this is likely to affect the ephemeral nature of subtitles as something which disappear as the frame of the scene changes. Unlike film viewing in the cinema, the *Receiver* of DVD will be able to see subtitles as many times as they wish and also process certain selected scenes, using DVD's interactive features. Videotape recording also made subtitles more permanent, but was not able to provide the interactive features in the way afforded by DVD. Such characteristics of DVDs give rise to several potentials. One is the emergence of DVD as an audiovisual database which allows the user to compare the original line in SL[1] against the subtitles and the dubbed version in a selected TL. This will provide subtitler training resources (Chen, 2003) and also background research resources for subtitlers to learn from prior solutions in terms of certain colloquial phrases, expressions, etc. which may be relevant to their assignment in hand. Given that a new medium always finds an ideal content, it is likely that as DVD becomes more widespread new content will be developed to take advantage of the unique features of this medium. One possible candidate may be animations which have recently seen significant advancement with digital technology.

As demonstrated by recent examples such as Web localisation, the demand for DVD localisation will be market-driven and new optimum procedures will ensue in response. While each new medium will demand different localisation procedures – making it difficult to envisage exact details – there may be some key aspects within today's localisation processes which remain as core to the emerging practices of *teletranslation*. The next section focuses on a particular dimension of localisation which the author considers as significant.

4. Significant dimension of localisation: Internationalisation

4.1 Today's internationalisation

The author maintains that localisation provides a theoretical basis for emerging language support. One particular dimension which seems to have a far-reaching impact is the concept of making allowance for *localisability* and *translatability* when developing the source content in SL. This is the process called internationalisation (see Esselink, 2000; Kano, 1995) increasingly applied to products which are subject to subsequent localisation.

Confusion over different definitions of the terms as regards to localisation, internationalisation and globalisation are often pointed out (Esselink, 2000). For the purpose of this paper, the author adopts the definition by Cadieux and Esselink (2002) who suggest: Globalisation = Internationalisation + N x Localisation. This formula shows that globalisation of a product or service involves internationalisation and localisation into a given number of locales. Both internationalisation and localisation therefore can be seen as key steps to achieve globalisation. In particular, internationalisation means preparatory tasks for subsequent localisation and therefore can be understood as localisation-enablement (Cadieux & Esselink, 2002). As such this approach makes a stark contrast with conventional translation which has typically taken the SL text as a given and thus applied to already completed SL text. In fact, one of the major characteristics of conventional translation has been the constraint imposed by SL text in relation to TL text production. By comparison, the internationalisation process aims to deal with foreseeable localisation and translation difficulties in advance at the inception of SL content.

The internationalisation process applied to software products involves technical adjustments to externalise all translatable components (Esselink, 2000). This means that internationalised products are designed and developed with such necessary modifications in mind as different character encoding systems (single byte vs. multiple byte characters), different lengths of TL text in relation to those of SL text, different conventions for expressing time, date and any other culturally specific elements in the *Receiver* context. Internationalisation of e-commerce Web sites may involve adjustments of non-textual elements, ranging from the design of the page, appropriateness of certain icons and images, to payment methods most suitable in the target culture in addition to consideration in terms of user inputs such as different currencies, digits for telephone numbers or the need or absence of postal zip code, etc.

Some may see internationalisation in association with earlier attempts at the use of controlled language applied to the source text. Controlled language is primarily intended to make the *Message* more machine-friendly by eliminating in advance those elements which are known to be problematic when the *Message* is destined for computer-based translation. Similarly, the concept of pre-editing of text for the use of MT is not new. However, the most extensive application of internationalisation is a novel approach in that the *Message* is controlled in its inception in terms of its techni-

cal, linguistic and cultural dimensions in relation to its *Receiver*. The internationalisation process can thus be seen as a comprehensive effort to make the *Message* amenable to the subsequent human- and machine-based translation processes and the use of controlled language can be regarded as a subset of the same attempt. As such, it is a clear contrast to the traditional treatment of translation as an isolated downstream activity where the *Sender* in SL has typically no regard for the *Receiver* in TL when creating the *Message*.

It is easy to demonstrate the advantage of internationalisation in the context of *TMC* with media such as computer software and the Web. For example, by recognising the fact that a software product originally produced in the USA will be marketed in China, Japan and Korea, technical allowances can be made to accommodate double-byte character sets required by these Asian languages as well as other culture-specific aspects in terms of the overall design of the software. If such consideration is not given at the beginning of the software development cycle and is not technically inbuilt, expensive re-development is likely to be necessary. The question of how the proactive approach such as internationalisation will impact on the overall *TMC* will be relevant in exploring further implications of internationalisation for emerging language support.

4.2 Future implications

Localisation processes have evolved under highly competitive commercial conditions and therefore cost-efficiency is a key factor. This focus is manifested in such tools as Translation Memory and workflow programs to shorten the time of localisation while maintaining the quality of the output. The need for internationalisation also arose to conduct a localisation task in a much more efficient way. Internationalisation has the potential to be extended to become a widely accepted general practice to make SL content amenable for *TMC* to the extent that most content used in electronic communications platforms are in effect internationalised. Given their potential global reach, such content will have a greater chance of being subjected to *TMC* without the *Sender* necessarily intending such an outcome. This means that the *Sender* with the internationalised *Message* only prepared in one language now has a better chance of getting the meaning across different languages and cultures in the event of some kind of language support being subsequently applied. This will further affect the role of the *Translator* and also the *Sender* in SL as their roles may increasingly be synchronised. In the context of localisation, a *Translator* may be working directly with the software developer as in a future scenario suggested by Esselink (1999) and in some cases the translation process may change into something close to synchronous sight translation turning out the TL content almost simultaneously with SL content production. Extended internationalisation may see the *Translator* in some cases becoming a designer of SL content in view of the potential *Receiver* in TL. For example, with DVD localisation, internationalisation may be applied to the audiovisual content to allow for the strategic application of subtitling and dubbing in combination. This will completely change the assumption of conventional screen translation with which the source content is not able to be modified. Such an approach may require the *Sender* and the *Translator* to work in collaboration.

In his future scenario for the localisation industry, Esselink (1999) proposes a database-driven dynamic model of localisation supported by a multilingual database as well as Translation Memory. In this model, the key is to reduce the time lapse between the content creation and its localisation by maximising the leverage of prior translations and by way of efficient workflow systems. Unlike MT-based real-time services, today's localisation is not a synchronous process and this remains as a challenge. Localisation seeks language support solutions by optimising the use of technology and, as a result, created productive tools specifically designed for localisation tasks. The increased applicability of localisation methodologies to a wider area of translation is seen in a widespread interest in Translation Memory systems. The concept of re-use of prior translations in a systematic way has an apparent appeal and a certain degree of applicability to a wide range of commercial translation. It is likely to impact the whole cycle of translation as Translation Memory evolves into a more generic tool and thus an integral part of the translation process as word processing is today. In this scenario, every translator needs to recognise the possibility of his or her translation reappearing at some future time somewhere. At the same time, this could mean that the translator does not have to translate the same sentence ever again. The concept of internationalisation and the approach based on efficient re-use of existing translations may significantly affect the translation process in future.

5. Conclusion

This paper endeavoured to demonstrate how localisation has become established in the modern translation business and is providing insight into emerging practices of language support. Web localisation and localisation of audiovisual content on DVDs were discussed as examples of existing and potential *teletranslation* practices using the *TMC* framework. The paper highlighted the internationalisation process as a significant dimension of new language support, suggesting its potentially far-reaching consequences. It seems justified to assume that *teletranslation* is developing, building on today's localisation rather than conventional translation. The author thus argues that localisation provides a theoretical basis to the future language support which is emerging from the nexus of language and technology particularly shaped by the society's shift to the infrastructure based on digital communications technology. This is creating new contexts of *TMC*. To this end, localisation research holds a significant key to understanding the future direction of translation.

Endnote

1. Some DVD titles will have subtitles in the Source Language as well as in the Target Language(s).

References:

Cadieux, P. & B. Esselink (2002), "GILT: Globalisation, Internationalisation, Localisation, Translation." *LISA Newsletter* 11, 1.5. Online at http://www.lisa.org/archive_domain/newsletters/2002/1.5/ (consulted 21.2.2003).

Chen, S.J. (2003). "Training the Next Generation of Subtitlers." *LISA Newsletter* 12, 1.5.

Esselink, B. (1999), "The End of Translation as We Know it". In: *Language International* 11(5).

Esselink, B. (2000), *A Practical Guide to Localisation*. Amsterdam, Philadelphia: John Benjamins Publishing.

Kano, N. (1995), *Developing International Software for Windows 95 and Windows NT (Microsoft Programming Series)*. Redmond, Wash: Microsoft Press.

Karamitroglou, F. (1999), Audiovisual Translation at the Dawn of the Digital Age: Prospects and Potentials. In: *Translation Journal* 3 (3). Online at www.accurapid.com/journal/09av.htm (consulted 22.03.2003)

Lookwood, R. (1998), "Language technologies and technical communication opportunities in FP5." In: *Computing & Control Engineering Journal* 9 (6), 253-56.

Lookwood, R. (1999), You snooze, you lose. In: *Language International* 11 (4).

Nida, E. A. & C. Taber (1969), *The Theory and Practice of Translation*. Leiden: E.J. Brill.

O'Hagan, M. (1996), *The Coming Industry of Teletranslation*. Clevedon: Multilingual Matters.

O'Hagan, M. & D. Ashworth (2002), *Translation-mediated Communication: Facing the challenges of globalisation and localisation*. Clevedon: Multilingual Matters.

Shannon, C.E. & W. Weaver (1949), *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Minako O'Hagan is a lecturer at Dublin City University, Ireland, where she teaches translation technology and Japanese. Her current research project investigates the applicability of CAT (Computer-aided Translation) tools to translating subtitles for DVD audiovisual content. She is the author of The Coming Industry of Tele-translation (1996), and co-author, with David Ashworth, of Translation-mediated Communication in the Digital World (2002).