ENHANCING FREE ON-LINE MACHINE TRANSLATION SERVICES

Federico Gaspari

Centre for Computational Linguistics
UMIST
PO Box 88
Manchester M60 1QD
United Kingdom

F.Gaspari@postgrad.umist.ac.uk

Abstract

This paper focuses on the use of free on-line machine translation (MT) services on the Internet and discusses how they could be enhanced. Some of the key challenges of using MT technology in the Internet environment are addressed within a wide framework, so as to emphasise the relevance of web usability, user-friendliness and good interaction design for the successful deployment of on-line MT services.

The paper suggests in conclusion that the performance of free web-based MT systems can be greatly enhanced by adopting a user-centred design that takes into account the specific requirements of on-line interaction, with a view to improving the perception that the general public of Internet users has of MT technology at large.

1. Introduction and background

A variety of on-line machine translation (MT) services are currently available free of charge for a wide range of language pairs on the Internet, and over the last few years popular search engines, browsable directories and portals have offered links to such services (cf. Yang & Lange 2003). Also, an increasing number of monolingual web-sites incorporate links to free on-line MT services, explicitly encouraging visitors to use them in order to obtain machine translations of their own web-pages into more familiar languages.

As a result, more and more Internet users are exposed to and take advantage of free on-line MT systems when they come across web-sites whose content is only available in languages unknown to them. Macklovitch (2001:27) summarises this situation as follows:

> Thanks to the spectacular growth and the pervasiveness of the World Wide Web, more people today have access to and are actually using MT than ever before. This democratisation of MT is very recent and […] the result of this MT-at-a-click phenomenon is undeniable: millions of people have actually experimented with machine translation now and so have at least some idea of what the current technology can and cannot do. […] The Internet has profoundly transformed the MT business.

Some recent reports suggest that free web-based MT services can in fact prove helpful in providing quick access to the general content of web-sites and on-line material in general, since many Internet users seem to be prepared to accept and read raw unedited MT output (see e.g. Flanagan 1996, Flanagan 1997, Hutchins 1999 and Yang & Lange 2003).

The rapid and steady growth in the use of free on-line MT services indicates that they have the potential to become very popular and successful resources on the Internet. However, web-based MT software in general seems to suffer from poor usability and inadequate interaction design, which makes it user-unfriendly and difficult to interact with.

This paper argues that this situation is more than likely to have a negative impact on the idea and impression that Internet users have of MT technology. The main problem addressed by the following discussion is how the performance of free on-line MT services can be improved, so that their users can experience a successful and satisfying interaction, regardless of the linguistic quality of the output produced by the MT software.

2. Origins and evolution of free on-line MT

The first free on-line MT service – Babelfish – was launched at the end of 1997 through the search engine Altavista. This innovative service was (and still is) powered by the core engine of the well-known MT system Systran, and initially offered 10 major European language pairs (cf. Yang & Lange 2003:191).

It is interesting to note that the number of language combinations covered by Babelfish has currently increased to 19 different pairs, which include for example Japanese, Chinese and Korean (in both directions with English) and Russian (available only as a source language in combination with English)[1].

Yang & Lange (2003) report on the experience of Babelfish, including information and figures about the actual use of the service based on the translation log files. Their data shows that there has been a constant growth in the demand for on-line translations over time: between two randomly chosen census days in June 1998 and November 1999 the number of translation requests submitted to Babelfish doubled (ibid.:203).

Users were encouraged to provide feedback on the service, and according to Yang & Lange (2003:196--198) their reactions ranged from enthusiastic to horrified. Over a five-month period at the beginning of 1998 only less than 5% of users giving feedback are said to disparage the service.

Overall these figures referred to the use of Babelfish suggest that free on-line MT systems can play a significant role on the Internet and that they attract a great deal of interest from the multilingual community of web-surfers.

Before the birth of Babelfish there had already been attempts to integrate machine translation solutions in Internet-based applications, but these were available only to a limited group of users. On-line MT was in fact originally offered in 1994 by CompuServe to its registered users in selected discussion forums (cf. Flanagan 1996).

After this pioneering experience, the launch in late 1997 of the free on-line MT service Babelfish marked the beginning of large-scale use of web-based MT without restrictions. Since then, in the last few years other MT vendors have offered their own web-based MT services with multiple language combinations free of charge over the Internet.

Following the example of Babelfish these MT companies that are competitors of Systran have developed their own free on-line MT services as a way to showcase their own commercial products, in order to make them known to potential purchasers through the Internet. In general, free Internet-based MT services are "lighter" versions of fully-fledged commercial MT systems that have been made available on-line for promotional purposes.

3. New approach to improve free on-line MT

What should be emphasised here is that these web-based MT services do not seem to take into account design and interaction requirements that would greatly enhance their use in the on-line environment. The Internet poses unprecedented demands for free automated translation services in real-time, since taking advantage of MT technology to translate entire web-pages on-line certainly is a challenging enterprise.

This transition calls for a careful and thorough re-consideration of general as well as more subtle issues concerning how free web-based MT can be successfully used on the Internet, by favouring user-friendly and user-centred design suited for on-line robust operation and easy interaction.

[1] This information has been taken from the web-site of Babelfish at the URL `http://world.altavista.com` and is correct as of 1 December 2003.

The author is convinced that a visible breakthrough in on-line MT technology cannot be achieved in the short term simply by improving the linguistic performance of the systems. As a result, this paper stresses the relevance of web usability principles and user-friendly interaction design for the enhancement of free Internet-based MT services, without taking into consideration the linguistic performance of the translation software at all.

To the best of the knowledge of the author, so far the numerous issues of web usability and interaction design raised by the on-line automatic translation of web-pages have not been considered by the MT-related industry and research community. This paper argues that it is vital to devote attention to these areas in order to significantly enhance the experience of Internet users who are exposed to free on-line MT technology.

Although it is recognised that it is necessary and desirable to improve the linguistic performance of on-line MT software (e.g. in terms of quality of the output), the author is convinced that significant progress in this respect will not be achieved in the next few years to come.

As a result, this paper suggests that other ways should be explored to enhance the user experience and interaction with on-line MT services regardless of their linguistic quality, but rather addressing the degree of usability and user-friendliness of their design. The rest of the discussion will focus on a practical approach to the enhancement of currently available on-line MT software that relies on basic web usability principles.

4. Web usability and free on-line MT services

At its simplest, implementing web usability principles into the design of an on-line application or Internet site acknowledges the importance of a user-centred approach, trying to make the interaction of the user with a web-site as smooth and successful as possible (see for instance Nielsen 2000, Krug 2000, Visciola 2000, Postai 2001, Nielsen & Tahir 2001 and Brink et al. 2002).

Over the last few years web usability has grown into a discipline of its own, and a number of principles and criteria to guide good web design have been developed. Usability issues play an increasingly important role in the design of Internet-based applications such as on-line shopping services, e-commerce and e-banking web-sites.

Similarly, key usability factors may have a crucial importance for the widespread acceptance and success of on-line MT technology on the Internet. Some examples will be given to show a variety of simple features and additional functions that could be incorporated into currently existing free web-based MT services to increase the degree of their usability and improve their interaction design.

4.1. User profiling and user modelling

Two areas that could be usefully explored to enhance free on-line MT services are those of user profiling and user modelling. This paper contends that simple techniques of this kind could be easily and successfully applied to the design of free web-based MT services in order to improve their usability. Simply put, user profiling and modelling aim at publishing and presenting information (and services) on the Internet in a way that suits a user best.

In an on-line application such as an e-commerce web-site this approach typically involves prioritising information, tasks and visual elements (e.g. icons, buttons, links, banners, etc.) that a particular user is most likely to be interested in. In other words, the user experience is tailored to the needs of individuals, and web-pages are adapted so as to give prominence to features and tasks that particular users are likely to be looking for, whereas the rest of the content (that may for instance be more interesting for other users) is kept in the background and receives less visibility.

What happens to people who access free on-line MT services is that all of them are presented with the same standardised web-page and graphic user interface over and over again. Every time they wish to submit a translation request to the service they need to specify the same information about the source and target languages they want.

User profiling and user modelling rely on information about the on-line behaviour and preferences of users that is gathered by means of cookies or agents, that follow and track the choices made by users while interacting with a web-site or on-line service. These user profiles can then help to determine the form and the sequence in which information or Internet content is presented to the users on the web-pages that they are looking at.

In the specific case of free on-line MT services, for example, user profiling and modelling could consist in identifying returning users in terms of language preference, prioritising the options that they are likely to choose. It seems in fact reasonable to assume that the large majority of stable or returning users of free web-based MT systems will most of the time choose the same source and target language combination when submitting a translation request.

One simple way of identifying and profiling returning users could be through registration, whereby stable users of free on-line MT services could be asked to create their own user account by specifying e.g. their mother tongue (i.e. the language they are most likely to translate into when they use the MT service), if they wish to do so. By relying on this user-specific information, every time registered users log into the service they would be offered only the options that are relevant and useful to them, e.g. in terms of available language pairs.

There is no space here to discuss in more depth the possibilities and technical details of how user profiling and modelling could be added to free web-based MT services. Rather, the intention is to briefly illustrate some simple features that would greatly enhance the degree of their usability, in favour of a user-centred interaction design.

What should be particularly emphasised in this respect is that at the moment no free on-line MT service provides any sort of customisation of its own graphic user interface to accommodate for the individual preferences and requirements of returning users.

4.2. Automatic source-language detection

Automatic source-language detection is another similar neat feature that free on-line MT services may adopt in order to increase the level of their user-friendliness and to enhance their usability.

The services could in fact help the user to automatically identify the language in which a web-page is written, and accordingly restrict the choice to the relevant language pairs, by showing only the combinations that are available with the source language concerned.

So if a web-page that a user wishes to translate is written in language A, the on-line MT service could automatically recognise this language, and only show on the screen the language combinations that are available with that source language, from which the user can then choose their preferred target language (i.e. A→B, A→C, A→D, etc., but not B→A, C→A or D→E).

In an informal survey the author has observed that at the moment the users of free on-line MT software are always shown all the possible options with all the language combinations that are available in the service. It is up to the users themselves to select the correct source and target language to obtain the translation of a web-page that they wish to read in a language different from the original.

This process is time-consuming and not particularly user-friendly, since users are always forced to view all the possible options and sift through them to identify which ones (if any) actually apply to the translation task they intend to perform.

If the language of the on-line input could be automatically detected by the MT service, the steps to access and start the rest of the procedure would become much simpler and quicker, since for instance the free MT service could dynamically adapt the on-line submission form. It could do so by showing only the relevant language combinations that it supports for the task at hand, i.e. the target languages that it can combine with the source language concerned.

In this scenario, users would only be expected to provide the free on-line MT engine with the URL of the web-page for which they want the translation, and the target language that suits them would be chosen from among those actually available. In summary, the initial customisation of the interaction procedure depending on the language of the source web document would save some effort to the users, guiding their steps towards relevant directions.

5. Enhancing free on-line MT services

The lack of flexibility and adaptation in the interaction and navigation design of currently available free on-line MT services seems to be a major shortcoming, since it may lead to difficulty of use, and as a result to a negative image of MT technology altogether. On the other hand, experience in a number of Internet-based applications (e.g. on-line shopping and e-commerce web-sites) has shown that taking into account web usability principles and good interaction design criteria can significantly increase user satisfaction, and ultimately benefit the productive use of the services.

A general tenet of web usability is that operations and actions required by on-line applications should be kept to a minimum or as much as possible hidden from the users, in order to benefit the interaction process. As a matter of fact, if users are expected to follow predefined procedures or to interact with the application on a step-by-step basis and by means of several explicit operations (as is currently the case in free on-line MT services), mistakes and problems can occur unless users are very experienced.

On the contrary, the more intuitive and straightforward the interaction is made, the more likely it is that people – including novice or unexperienced users – complete it successfully, without making mistakes at any stage of the process. It is argued that such a usability-oriented approach, briefly exemplified in practice by the simple features and design guidelines suggested above, could be successfully applied with very limited investment to free on-line MT services in the near future.

6. Future developments for on-line MT

The problem that has been discussed so far is how to improve the performance of free web-based MT services so that their users experience a successful interaction, regardless of the linguistic quality of the output produced by the MT software.

In the view of the author, this new approach to the enhancement of on-line MT technology may have a number of interesting future developments, some of which will be briefly reviewed here.

The methodology proposed in this paper to improve free web-based MT services relies on the cross-fertilisation between on-line MT on the one hand and a number disciplines that are concerned with the investigation of Internet user behaviour on the other. This approach is driven by insights gained from areas as diverse as HCI (Human-Computer Interaction), ergonomics, cognitive psychology, web usability and web localisation, and attempts to apply them to on-line MT.

The primary objective of this new method is to explore how the user experience with free web-based MT services can be enhanced, by taking into account the specifics of on-line use and interaction. As has been illustrated above, this can be achieved by applying to free Internet-based MT principles and techniques that have recently emerged in other areas of the Internet environment.

One of the short-term aims of this work, then, is the application to the design of on-line MT services of web-usability principles that have emerged for the improvement of e.g. e-commerce and on-line shopping web-sites. The author is convinced that pursuing this approach would yield visible results in the near future for the enhancement of free on-line MT, making it a (more) successful tool and a valuable resource for Internet users.

Medium-term developments of this usability-oriented approach range from the establishment of benchmarks for the evaluation of free on-line MT services to the formulation of guidelines for the successful integration of web-based MT solutions into monolingual web-sites for multilingual dissemination purposes.

Finally, a related area for future action which may represent a potential long-term spin-off from this research approach includes the development of web-authoring tools, design templates and platforms to create monolingual web documents that are amenable to free on-line MT services. This last point seems particularly interesting, since it is likely to attract considerable interest, and to the best of the knowledge of the author so far no action along these lines has been taken.

The brief discussion that follows raises some key points that concern the extent to which the web-pages of monolingual web-sites are suitable for free on-line MT services in order to obtain machine-translated multilingual versions in other target languages. The suggestions put forward here are based on the author's contention that some new support services and tools will be needed to expand and improve the current applications of free on-line MT. Given the current popularity of web-based MT services, the need for them may be closer than one would expect.

One example may be the development of tools and resources to prepare on-line MT-friendly web-pages and web-sites, for instance by creating pre-defined templates that web-designers could adopt to present the contents of their own web-sites. These templates could have authoring guidelines and MT-friendliness checking functions at the linguistic level, similar to those that are commonly found nowadays in spellcheckers, thesauri and grammar checking facilities of word-processors. As a result, the creation of such tools would not require substantial investments from scratch, but would seem to have huge business potential in connection with the development of free on-line MT services.

At a very simple level, this linguistic checklist would show to the web-designer that the web-site they are creating contains texts with long sentences, thus recommending that their length is reduced. At a more advanced level, these functions would help for instance to detect single hot-words that web-designers would like to use in hyperlinks, in case they have different meanings in the source language, and as a result they may have different translations into another language.

In such cases of polysemy, then, on-line MT services may not render the word correctly in the target language(s), giving rise to wrong translations, misunderstandings and misleading choices in the target language. By relying on the help of this MT-aware authoring tool, web-designers would be advised to avoid the use of such potentially ambiguous words, and may for instance opt to replace the hot-word concerned with another unambiguous word or even a language-neutral icon. In doing so, they would refer to the same concept, function or operation by means of a visual language-independent message, which would be kept in the machine-translated versions of the document as well.

The author believes that the examples that have been shortly illustrated here provide a reasonable picture of tools that may be necessary in the future in connection with free on-line MT services offered in real time on the Internet. This final part of the discussion has aimed at providing some speculations on possible future developments, and it is hoped that continuous research is carried out in order to monitor the trends and opportunities that arise in the exciting field of on-line MT.

7. Conclusion

Free on-line MT services have been available on the Internet for a few years now, and they are powerful tools that attract growing interest. Even though MT technology is easily accessible on the Internet today, in general free on-line MT services are not easy to use and interact with successfully. This paper has addressed the need to enhance the performance of these services, with a view to improving the perception that the general public of Internet users has of MT technology.

Usability and user-friendliness issues are currently receiving growing attention in the Internet community, mainly in connection with the development of e-commerce and web-sites that specialise in on-line business transactions. A similar direction seems the best way to go to raise the profile of free on-line MT in the near future, so that it may gain credibility among a larger population of Internet users.

Serious work is needed in this area, so that lessons learned in other on-line domains can be applied to web-based MT. Looking at the current picture of free on-line MT services, it seems that some of the basic guidelines that enhance usability and user-friendliness could be effectively implemented with limited effort in the short term, so as to become standard requirements for high-profile Internet-based MT providers.

The paper has mentioned some areas for future action, identifying some possible developments that are either needed or foreseeable for the future. In conclusion, since on-line MT is a young and innovative technology, it seems that most of its potential for further development has not been revealed yet.

References

Brink, T., D. Gergle & S.D. Wood (2002). *Designing web sites that work. Usability for the Web*. Morgan Kaufmann.

Flanagan, M. (1996). "Two years online: experiences, challenges and trends". In *Expanding MT Horizons: proceedings of the Second Conference of the Association for Machine Translation in the Americas, 2-5 October 1996, Montreal, Quebec, Canada*. 192-197.

Flanagan, M. (1997). "MT Today: Emerging Roles, New Successes". *Machine Translation*. 12:25-27.

Hutchins, J. (1999). "The development and use of machine translation systems and computer-based translation tools". In *Proceedings of the International Conference on Machine Translation & Computer Language Information Processing*, *26-28 June 1999, Beijing, China*. C. Zhaoxiong (ed.). 1-16. [Available on-line at the URL `http://ourworld.compuserve.com/homepages/WJHutchins/Beijing.htm` - Accessed 1 December 2003].

Krug, S. (2000). *Don't Make Me Think: A Common Sense Approach to Web Usability*. New Riders.

Macklovitch, E. (2001). "Recent Trends in Translation Technology". In *Proceedings of the 2nd International Conference The Translation Industry Today: Multilingual Documentation, Technology, Market. 26-28 October 2001, Bologna, Italy*. 23-47.

Nielsen, J. (2000). *Designing Web Usability: The Practice of Simplicity*. New Riders.

Nielsen, J. & M. Tahir (2001). *Homepage Usability: 50 Websites Deconstructed*. New Riders.

Postai, S. (2001). *Siti che funzionano. Quando Web design non significa disegni su web*. Hops Libri.

Somers, H. (ed.) (2003). *Computers and Translation: A translator's guide*. John Benjamins.

Visciola M. (2000). *Usabilità dei siti web*. Apogeo.

Yang, J. & E. Lange (2003). "Going Live on the internet". In H. Somers (ed.). 191-210.