

Chapter 4 Feature checklist applied – a comprehensive comparison of MAHT packages.

4.1 Feature checklist.

	Trados	Déjà Vu	SLDX	Wordfast
1.1.1	Y	Y	Y	Y*
a	Y	Y	Y	Y
b	user specified: end rules, skip rules, source and target segmentation rules; optionally: structure recognition (styles, paragraphs, font sizes); user specified importance of: tags, numbers, formatting information, term lists.	user specified sets of splitting rules and exceptions to them, (one for the source language and another for the target language)	user specified language specific segmentation end rules and exceptions to them	user specified segmentation rules (end of segment punctuation) set for both source and target languages
c	-	-	-	-
d	Y	Y	Y	Y
e	-	-	-	-
1.1.2	Y	Y	Y	Y
a	N the whole text	N the whole text	N the whole text	N the whole text
b	Y	Y	Y	Y
c	-	-	-	-
1.1.3*	(212 source and 215 target segments/ 188 correct pairs) 87,44% of correct alignments	**(208 source and 205 target segments/ 37 correct pairs) 18,05% of correct alignments	*** (112 source and 107 target segments/ 36 correct pairs) 33,64% correct alignments (102 source and 107 target segments/29 correct pairs) 27,10% correct alignments	****(212 source and 213 target segments/ 36 correct pairs) 16,9% of correct alignments
1.1.4				
a	Y	Y	Y	Y
b	Y	Y	Y	Y

c	Y	Y	Y	Y
d	-	-	-	-
1.1.5				
a	Y	Y	Y	Y
b	Y	Y	Y	Y
1.2.1				
a	Y	Y	Y	Y
b	Y	Y	Y	Y
c	N	N	N	N
1.2.2				
a	Y	Y	Y	Y
b	Y	Y	Y	Y
c	N	N	N	N
1.2.3	Y	Y	Y	Y
1.2.4				
a	option	Y	Y	option
b	option	N	N	option
c	option	N	N	option
d	-	-	-	-
1.3.1	Y	Y	Y	Y
a	Y	Y	Y	Y
b	Y	Y	Y	Y
1.3.2				
a	option	N	option	option
b	option	Y	option	option
c	option	N	N	option
d	option	-	-	option
1.3.3	1	1	1	1
1.4.1	Y	Y	Y	Y
a	Y	Y	Y	Y
b	Y	Y	Y	Y
1.4.2	Y	Y	Y	Y
1.4.3	Y	Y	Y	Y
a	Y	Y	Y	Y
b	Y	Y	Y	Y
c	Y	Y	Y	Y
d	Y	Y	Y	Y
e	export of data selected according to TM specific attributes such as: source and target language text, date of creation, author, editor, other attributes (e.g. status, client domain)	export of data according to the following TM attributes: source and target language, client, subject.	export of TUs according to TM attributes such as: date of creation, author, date of last modification, author of last modification, date of last use, frequency of use. All attributes can have a user-	by deletion of selected TUs according to: target and source language, target and source text, usage count, creation date, up to 4 user specified attributes (e.g. author of TM)

			specified value.	
f	Y combination of attributes, logical rules AND, OR , NOT	Y combination of attributes: operations (less than, less than or equal to, equal to, greater than, greater than or equal to, like, not like) logical rules AND, OR.	Y combination of attributes: operations (equal to, not equal to, less than, greater than, less than or equal to, greater than or equal to), AND, OR rules.	Y (combination of attributes from e) according to logical rules: and/or, less than,greater than, equal to, includes, includes not)
1.4.4	Y	Y	Y	Y*
1.4.5	Y	Y	Y	Y
1.4.6	N	N	N	N
1.4.7				
a	option	N	option	option
b	option	Y	option	option
c	option	N	N	option
d	-	-	-	-
1.4.8	Y	Y	Y	Y
1.5				
a	Y	Y	Y	Y
b	Y (email)	Y(email)	Y (email)	Y(email)
2.1				
2.1.1	*	**	***	
a	0	*4 segments, 18 words	3 segments, 12 words	0
b	2 segments, 46 words, approx. 2% of the text length	b+c 60 segments out of 119 total sentences	b+c 5 segments, 71 words approx 3% of the text length (2323 words,163 segments)	b+c* 6 segments, 175 words, approx 7% of the text length
c	1 segment, 35 words, approx. 2% of the text length	-	-	-
2.1.2	Y	N	Y	Y
2.2.1				
a	Y	Y	Y	Y
b	Y	Y	Y	Y
c	Y	Y	Y	Y
d	N	N	N	N
2.2.2	Y	Y	Y	Y
2.2.3	Y	Y	Y	Y
a	Y	Y	Y	Y
b	Y	Y	Y	Y
c	Y	Y	Y	Y
d	Y	Y	Y	Y
e	Y	Y	Y	Y
2.2.4				

a	Y (user set formatting penalty)	N	Y (user set)	Y (user set percentage penalty)
b	Y(user set formatting penalty)	N	Y (user set)	Y (user set percentage penalty)
c	Y(user set formatting penalty)	N	Y (user set)	Y (user set percentage penalty)
d	user set	5	Y (user set)	user set (as in a,b,c)
e	Y	Y	Y	Y
f	Y	Y	Y	Y
*g	Y	Y	Y	Y
h	N	N	N	Y
2.2.5				
a	Y	Y	Y	Y
b	Y	Y	Y	N
c	Y	Y	Y	Y
d	Y	Y	Y	Y
2.2.6	Y	Y	Y	Y
2.3				
a	Microsoft Word, TRADOS T-Window, TRADOS TagEditor.	Déjà Vu Interactive	program's own SDL Edit module	Microsoft Word
b	Microsoft Word	Déjà Vu Interactive	SDL Edit	Microsoft Word
c	Multiterm (TMS)	Terminology maintenance (TMS)	TermBase (TMS)	+Tools
3.1	Y	Y	Y	Y
3.2				
a	Y	Y	Y	Y
b	Y	Y	Y	Y
3.3	Y	Y	Y	Y
3.4	customizable windows (as in standard Windows applications), standard layout suggested.	customizable window sizes and positions (vertical vs horizontal)	customizable window sizes and positions (vertical vs horizontal)	customization options, program's own suggested settings.
4				
a	Y	Y	Y	Y
b	Y	Y	Y	Y
c	Y	Y	Y	Y
d	Y	Y	Y	Y
e	Y	Y	Y	Y

f	Y	Y	Y	Y
5				
5.1	Windows 98/Me/NT (SP3 or higher), Windows 2000/XP. Windows XP or Windows 2000 recommended.	Windows 98/Me/NT/2000XP	Windows 98/NT 4.0/ 2000/XP	Windows 95/98/Me/NT4/2000/XP Mac OS 7, 8, 9.xx, Mac OSX Linux (Mandrake, XFCE, VMWare & Office 97)) Linux (Suse, KDE2 + Crossover & Word 2000)
5.2				
5.2.1	PC with Pentium or compatible processor/ Pentium II or higher	recommended: Pentium III 600MHz	Pentium 90/ Pentium 350 or higher	no info provided (it may be safely said that system requirements are low, as Wordfast is only a Word macro)
5.2.2	64 MB RAM (128 MB recommended) on Windows 98/Windows Me and 128 MB RAM (256 MB recommended) on Windows NT/Windows 2000/Windows XP.	recommended: 256MB	64MB/ 256MB	no info provided (it may be safely said that system requirements are low, as Wordfast is only a Microsoft Word macro)
5.2.3	(separate for every component of the package)	no info provided	50MB (to install)/ more	no info provided (it may be safely said that system requirements are low, as Wordfast is only a Microsoft Word macro)
6	5	4	4	2
6.1	user guides (very detailed), tutorial demos, FAQs.	extensive downloadable materials:a manual, tutorial; a FAQ list. Website with a comprehensive collection of links	extensive downloadable materials: introductory guides, getting started documents, how-to docs, manuals. FAQ	downloadable manual

		to various reference documents, tool reviews etc. (http://www.transref.org/)	lists	
6.2	fee-based consulting services, open seminars, on-site customized training courses	fee-based training programs, both on-site and at software developer's facilities available in several countries organised on an individual basis	onsite training for an additional fee, free QuickStart online training via Webex conference facilities (for some program editions); fee-based consulting services (seminars, training etc.)	none
6.3	searchable on-line Knowledge Base (with articles, tips and tricks etc), support service based on individual contracts, customer FAQs, newsletter, network of offices. http://www.translationzone.com/ -a separate website for freelance translators with: first steps guides, user guides, FAQs, knowledge base, user support schemes	on-line support center with: mailing lists, newsgroup, a searchable Knowledge Base containing FAQs, how-to docs etc., network of offices	on-line fora (discussion groups), Newsletter, Whitepapers-Articles page, on-line support page, network of offices	developer's website, discussion groups (very active), author's email
*7	695 euro	990 euro	995 euro	180/90 euro**
8	Y	N	Y	Y

4.2 Comments

*1.1.1 as Wordfast is not a “proper” software package, it is simply a Word template, (it could be said to be a bit “buggy” or experimental) and making it work requires some determination on the part of the user. Fortunately all the required steps are described in the manual.

*1.1.3 Obviously the percentage of correct alignments heavily depends on a number of variables, such as: type of text, its structure, segmentation settings etc. To provide some basis for comparison the following procedure has been employed: segmentation settings were set on default, an 18-page-long section of the EU Accession Treaty (<http://www3.ukie.gov.pl/dokumenty/ta/aa00002.en03.pdf>) was segmented and aligned with its translation provided by UKIE (<http://www3.ukie.gov.pl/dokumenty/tapl/aa00002.pl03.pdf>), the percentage of correct alignments was calculated (an alignment requiring no editing was considered correct). As *.pdf files cannot be directly processed by CAT tools, text from those files was extracted with the use of Ghost View 4.3 and saved as *.doc files.

**1.1.3 Due to demo version limitations the text had to be split into two smaller texts, hence two percentage values are presented.

***/** 1.1.3 (it has to be noted that minimal editing, that is, joining of several source or target segments at least tripled the number of correct pairs)

*1.4.4 Wordfast stores TMs in a tab delimited txt format. Therefore, a SL segment may be simply copied and pasted to another TM (the user needs to be careful to preserve the format of a TM and to paste the text between the tabs).

*2.1.1 Results obtained with the use of the Analyse function of TRADOS Translator's

Workbench

**2.1.1 Results obtained with the pretranslate function, with the "fuzziness" option set on maximum (9). This option determines how fuzzy a match should be for the program to accept it as a match and include it in the pretranslation. It has to be taken into account that some matches found by the program may fall below the 50% threshold set for the other packages. This explains the relatively high number of fuzzy matches.

*2.1.1b Results obtained with the use of the Analyse function of Wordfast, which presents a single value for fuzzy matches (50-94%). Therefore the result shown here is a sum total of b and c values.

*2.2.4 g) It is commonly suggested by CAT software providers that the useful threshold set for fuzzy matches should not be lower than 75%.

*7 – All prices are provided here solely for the purpose of providing a general idea of CAT software costs. It has to be taken into account that prices are varied according to the contents of a package, the type of customer and are subject to promotional offers. The author decided to list prices of freelance editions (as of April 12, 2003) given on software provider's official websites, disregarding all discounts. Therefore, to acquire a more detailed view of prices the reader should visit the following websites: for Trados <http://www.trados.com>, for Déjà Vu: <http://www.atril.com>, for SDLX: <http://www.sdlintl.com> , for Wordfast: <http://www.champollion.net/>

7** 180 euro for customers from: The USA, Mexico, Canada. Andorra, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Liechtenstein, Luxembourg, Monaco, Portugal, Spain, The Netherlands, Norway, Sweden, Switzerland, The

United Kingdom. Australia, Brunei, Hong-Kong, Japan, Korea, New Zealand, Singapore, Taiwan. Bahrein, Israel, Kuwait, Qatar, Saudi Arabia, UAE. (90 euro for customers from countries other than those listed above, including Poland)

Conclusions

In conclusion, the author can safely confirm the high usefulness of MAHT packages which are especially tailored to the needs of translators and greatly enhance their productivity, which is confirmed by the fact that all tools under scrutiny meet almost all requirements contained in the feature checklist.

Furthermore, in the wake of this sample study, the author may draw the conclusion that all CAT packages provide very similar functionalities, although they achieve the end not necessarily by exactly the same means.

As regards the results of the comparative evaluation performed with the use of a checklist proposed by EAGLES, the author may conclude that the checklist in its present form, despite its limited scope, does yield useful and quantifiable evidence, especially as regards the performance of the TM and alignment module.

What is more, the study shows how important the expertise of the user is in view of the fact that CAT programs enable the translator to micro manage the program's settings and that using proper segmentation rules for particular texts has a tremendous influence on the tools' performance.

However, it has to be admitted that, although the feature checklist had been enhanced according to its authors' suggestions, it cannot fully account for the more complex and in-depth differences among the evaluated software packages. This result is a testimony to the fact that MAHT evaluation still remains somewhat of an uncharted territory.

Last but not least, the author can only hope that the use of CAT tools will become more commonplace in Poland. After all, as Danilo Nogueira, a CAT software reviewer put it (2002) "Translation memory software is software for translators. If you translate using Word for Windows (for writers) or, say, PageMaker (intended for DTP work) you are using knives to tighten screws."

References

ALPAC. 1966. *Language and machines: computers in translation and linguistics*. Washington, D.C.: A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council. Publication 1416.

Arnold, Doug, R. Lee Humphreys and Louisa Sadler (eds.) 1993. *Machine Translation, Special Issue on Evaluation of MT Systems. Volume 8, Nos. 1-2*. Dordrecht: Kluwer Academic Publishers.

Arnold, Doug, Lorna Balkan, Siety Meijer, R. Lee Humphreys, Louisa Sadler. 1994. *Machine Translation: An Introductory Guide*. <http://clwww.essex.ac.uk/~doug/book/book.html> date accessed: July 24, 2002

Benis, Michael. 1998. "Review of Atril's Déjà Vu 2. The Happy Hoarder." In: *Translation Journal. No. 1, Volume 2, January 1998*. <http://accurapid.com/journal/03TM1.htm> date accessed: June 3, 2003

Benis, Michael. 2002. *How the memory measured up* <http://transref.org/> date accessed: June 3, 2003

Briggs, Brian *The Importance of TMX for Translation Tool Buyers* <http://www.languagepartners.com/reference-center/whitepapers/l10nwp/tmx-importance.htm> date accessed: July 11, 2002

Brungs, Bettina. 1996. *Translation Memories als Komponente Integrierter Übersetzungssysteme*. Saarbrücker Studien zu Sprachdatenverarbeitung und Übersetzen, Band 7.

Budin, Gerhard. 1998. *Evaluierung von Translationleistungen: Machinelle Übersetzungen*. Courtesy of the author.

Crelin, Jonathan; Thomas Horn and Jenny Preece. 1990. "Evaluating Evaluation: A Case Study of the Use of Novel and Conventional Evaluation Techniques in a Small Company." In: Diaper et al. (eds.), 1990. 329-335.

Diaper, D., G. Cockton, D. Gilmore, B. Shackel (eds.) 1990. *Human Computer Interaction – INTERACT 90*. Amsterdam / New York / Oxford / Tokyo: Elsevier Science Publishers B. V. (North-Holland).

EAGLES. 1995. *Evaluation of Natural Language Processing Systems, EAGLES document EAG-EWG-PR.2. Version of September 1995* <http://www.ilc.pi.cnt.it/EAGLES> date accessed: August 11, 2002

EAGLES. 1999. *EAGLES Evaluation Working Group. Final Report. EAG-II-EWG-PR.1. Draft – March 1999*. <http://www.ilc.pi.cnt.it/EAGLES> date accessed: August 23, 2002

- Falcone, Suzanne. 1998. "Translation Aid Software. Four Translation Memory Programs Reviewed." In: *Translation Journal. No. 1, Volume 2, January 1998* <http://accurapid.com/journal/03TM2.htm> date accessed: June 3, 2001
- Falcone, Suzanne. 2000. "More Translation Memory Tools." In: *Translation Journal. No. 2, Volume 4, April 2000.* <http://www.accurapid.com/journal/12TM.htm> date accessed: June 3, 2003
- Feder, Marcin 2001. *Computer Assisted Translation. A Proposal for Tool Evaluation Methodology*. School of English, Adam Mickiewicz University, an unpublished Ph.D. thesis. Courtesy of the author.
- Feder, Marcin 2002a "Localisation in Poland. An Overview of The Polish Localisation Industry" in *TermNet News TNN 2002 74*
- Feder, Marcin 2002b. *A Tentative Proposal for Machine Assisted Human Translation (MAHT) Tool-Specific General Text Typology* Courtesy of the author.
- Heyn, Matthias. 1995. "Key Technologies. Impacts of Neural Network Technology on Computer Aided Translation". In: *TAMA'94 Proceedings*. 1995. 71-84.
- Hüberger, Andreas. 2002. *What You Need to Know About Translation Memories* <http://www.multilingualwebmaster.com/library/trmemories.html> date accessed: June 3, 2003
- Hutchins, John. 1996. "ALPAC: the (in)famous report" In: *MT News International* 14, June 1996, pp. 9-12. <http://ourworld.compuserve.com/homepages/WJHutchins/Alpac.htm> date accessed: June 3, 2003
- Kay, Martin. 1997. "The Proper Place of Men and Machines in Language Translation." In: *Machine Translation 12*. 3-23. (First published as a Xerox PARC Working Paper in 1980).
- Kingscott, Geoffrey. 1999. 'New strategic direction for TRADOS' In: *International Journal for Language and Documentation* November 1999
- Melby, Alan 1982. "Multi-Level Translation Aids in a Distributed System." *COLING 1982*.
- Melby, Alan 1990. *The Translator Workstation*. Department of Linguistics, BYU, Provo, Utah 84602 USA
- Nogueira, Danilo. 2002. "Translation Tools Today: a Personal View." In: *Translation Journal. No. 1, Volume 6, January 2002.* <http://www.accurapid.com/journal/19tm.htm> date accessed: June 3, 2003
- Pooley, David. 2002 *The importance of TMX* http://www.lisa.org/archive_domain/newsletters/2002/3.6/pooley.htm date accessed: February 22, 2003

Ray, Deborah S., Eric J. 1999 "Good, Fast, Cheap: Translation Memory Systems Offer the Potential for All Three" *Technical Communication the journal of the Society for Technical Communication* May 1999 (V. 46, No. 2).

Schüller, Thilo. 1995. *Integrierte Übersetzungssysteme*. Saarbrücker Studien zur Sprachdatenverarbeitung und Übersetzen, Band 1.

Somers, Harold L. 1997. "Machine Translation and Minority Languages." *Translating and the Computer* 19. 1997.

SDL International *Getting Started with SDLX*
<http://www.sdlintl.com/products/products-and-services/downloads.htm> date accessed: June 3, 2003

Spies, Christina. 1995. *Vergleichende Untersuchung von Integrierten Übersetzungssystemen mit Translation-Memory-Komponente*. Saarbrücker Studien zu Sprachdatenverarbeitung und Übersetzen Band 3.

TRADOS Corporation 2002 *Translator's Workbench User Guide*
<http://www.translationzone.com/download.asp> date accessed: September 8, 2002

Trujillo, Arturo 1999. *Translation Engines: Techniques for Machine Translation*. London: Springer.

Vasconcellos, Muriel (ed.). 1988a. *Technology as Translation Strategy. American Translators Association Scholarly Monograph Series, Volume II*. State University of New York at Binghamton (SUNY).

Vasconcellos, Muriel. 1988b. "Factors in the Evaluation of MT: Formal vs. Functional Approaches". In: Vasconcellos (ed.) 1988a. 203-213.

Webb, Lynn E. 1999. *Advantages and Disadvantages of Translation Memory: A Cost/Benefit Analysis*. An unpublished M.A. thesis. Monterey: Monterey Institute of International Studies.
<http://www.webbsnet.com/translation/thesis.html> date accessed: June 3, 2003