



eWin: European Waste Information Network Workpackage 4.6.1 - Translation System

CREDO
University of Sheffield
United Kingdom

Description

Currently available natural language translation systems will be evaluated. The best candidate will be selected and integrated into the generic system.

Abstract

For the purposes of the EWIN project it is necessary to evaluate Machine Translation tools running in the widely used Operating Systems and Web Browsers. Machine Translation (MT) is the term used for software specialised in translation. MT software incorporates linguistically sophisticated technologies from dictionaries to a collection of linguistic rules that make it possible to provide a good quality translation without using humans. The software is to be used in the Client side of the architecture to translate the data returned from the EWIN server.

Keywords: Machine Translation, Information Gateways

Version details

<i>Version:</i>	2
<i>Date:</i>	24-11-2000
<i>Circulation:</i>	Consortium only
<i>Status:</i>	submission

Document History

Version	Date	Author / Responsible	Description
1	17-11-2000	CREDO, Dr. Ilias Stergiou	Draft
2	24-11-2000	CREDO, Dr. Ilias Stergiou, M. Kraner	submitted

1. Table of Contents

1. <i>Table of Contents</i>	2
2. <i>Introduction</i>	3
3. <i>Machine Translation</i>	3
3.1. <i>Transparent Language</i>	4
3.2. <i>SYSTRAN</i>	5
3.3. <i>Lernout & Hauspie's - AppTek</i>	6
3.4. <i>LTSS</i>	6
4. <i>Machine Translation products</i>	7
4.1. <i>L&H iTranslator™ Enterprise</i>	7
4.2. <i>L&H iTranslator Software Developer Kit (SDK)</i>	7
4.3. <i>L&H iTranslator Express</i>	7
4.4. <i>L&H Power Translator Pro 7</i>	8
4.5. <i>L&H Simply Translating</i>	9
4.6. <i>L&H WebTranslator</i>	9
4.7. <i>Transparent Language - Enterprise Translation Server</i>	9
4.8. <i>Transparent Language - TranscendRT SDK</i>	10
4.9. <i>Transparent Language - Easy Translator</i>	10
4.10. <i>Transparent Language - Desktop Translator</i>	10
4.11. <i>Transparent Language - Transcend</i>	11
4.12. <i>SYSTRAN Personal</i>	12
4.13. <i>SYSTRANET</i>	12
4.14. <i>SYSTRAN PROfessional Standard</i>	12
4.15. <i>SYSTRAN PROfessional Premium</i>	13
4.16. <i>SYSTRAN Enterprise</i>	13
5. <i>Machine Translation Product comparison</i>	14
6. <i>Web-based translation services</i>	15
6.1. <i>FreeTranslation</i>	15
6.2. <i>AltaVista</i>	16
6.3. <i>Systranet</i>	16
6.4. <i>e-lingo</i>	17
6.5. <i>L&H iTranslator™ Online</i>	17
6.6. <i>Globalink</i>	17
6.7. <i>Forthnet</i>	17
7. <i>References</i>	17

2. Introduction

In the EWIN project there are partners from 6 different countries (Greece, Great Britain, Germany, Spain, Norway and Denmark), speaking 6 different languages. The content of the repository is usually on the language of the country of origin and is not translated in the other 5 languages. Therefore, the repository hosts information in 6 different languages making it difficult for anyone who is not a polyglot to utilise the information provided.

The target market of EWIN is any Internet user interested in waste. The EWIN server, while available on a global level, underscores the serious limitations in our ability to communicate across language barriers. Today, taking English as an example, less than 54% of all Internet users are native-speakers of English and, by 2005 this number will drop an additional 10%, to under 43% (Figure 1). In the next 6 years, the Internet will experience a 150% increase in usage among non-English speakers [CE, 99]. Clearly, the power of the EWIN service provided is limited to our ability to communicate successfully to our target market.

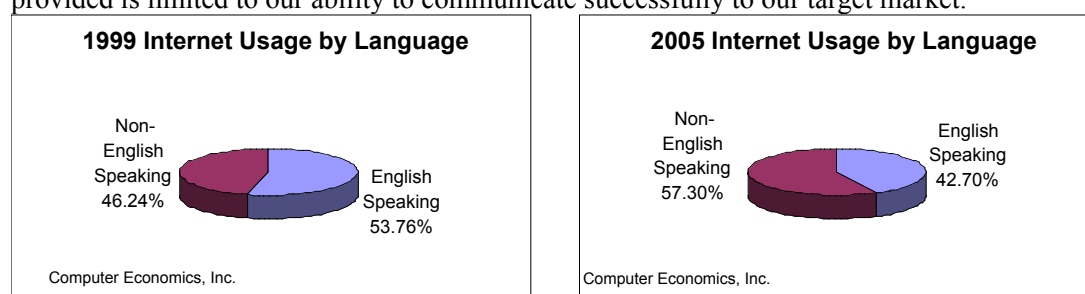


Figure 1

This deliverable in the third Chapter describes what Machine Translation is, what categories exist and the engines available. The fourth Chapter describes the products available. Chapter five compares the products and proposes one, having in mind a “waste” consultant as the target user. Finally Chapter six presents the Web based translation engines that a “waste” consultant can use for free.

3. Machine Translation

Machine Translation definition

Machine Translation (MT) is the term used for software specialised in translation. MT software incorporates linguistically sophisticated technologies from dictionaries to a collection of linguistic rules that make it possible to provide a good quality translation without using humans.

The process of MT involves breaking down complex and varying sentence structures, identifying parts of speech, resolving ambiguities, and synthesising the information into the components and structure of the new language.

Machine Translation categories

There are three MT software categories:

- Direct** It is a simple word by word translation without interpreting sentence structure. The result is a translation that is in the word order of the source text, which can result in a significant loss of meaning and readability. The comprehension and translation quality is often unacceptable.
- Transfer** Performs semantic analysis of the source language sentences and then transfers that information into the target language based on a set of rules specific to that language direction. The quality of the resulting translations is significantly higher than the simple Direct translations.

Interlingua In theory, this method would input the source language into an intermediate repository where expressions, sentence semantics and forms of expression would be replaced, independent of any language, and then the intermediate language, or meta-language, would be translated into the target language. This theoretical method has yet to be used with any commercial applications.

Products of the Direct category are of very poor quality and will not be studied in this document. The Interlingua category theoretically offers the best quality results but no commercial product is currently available. For the purpose of this document the Transfer category Machine Translators will be studied.

Machine Translation engines

Machine Translation software requires a huge amount of research and development. A recent major study of translation technology by an independent research firm confirmed that a new first-class Machine Translation engine would require 200 to 300 person-years to develop [Transparent, 00]. Today there are only a few companies that have produced Transfer category MT engines, as presented in the following sections.

3.1. Transparent Language

TranscendRT is Transparent Language's (<http://www.transparent.com/>) powerful, proprietary machine translation engine to rapidly translate text from one language to another using the Transfer method.

The translator [Transparent, 00] consists of the following components (as shown Figure 2):

1. **Segmenting:** The source document is divided into paragraphs, then sentences and finally words, all the while retaining formatting information. A sequence of filtering and scripting operations makes this process very flexible within the TranscendRT system.
2. **Morphological analysis:** The individual words are looked up in the dictionaries and phrases are identified. The TranscendRT has the ability to recognise a word in various forms; for example, identifying that "talking" is the gerund form of "talk."
3. **Functional analysis:** The TranscendRT determines how each word functions in the sentence and determines the part of speech for words that can function as more than one part of speech. The TranscendRT also determines which words are functioning as part of a recognised phrase.
4. **Syntactic analysis:** The TranscendRT determines the grammatical structure of each sentence, aided by information about the meaning of the words in the sentence.
5. **Transfer:** The elements of the sentence are reordered, inserted or deleted to form the syntax of the new language.
6. **Generation:** Creating the final sentence in the new language is a multi-step process that, in effect, reverses the steps for analysing the source sentence. When the process is complete, the TranscendRT restores the original formatting information as it applies to the new text.

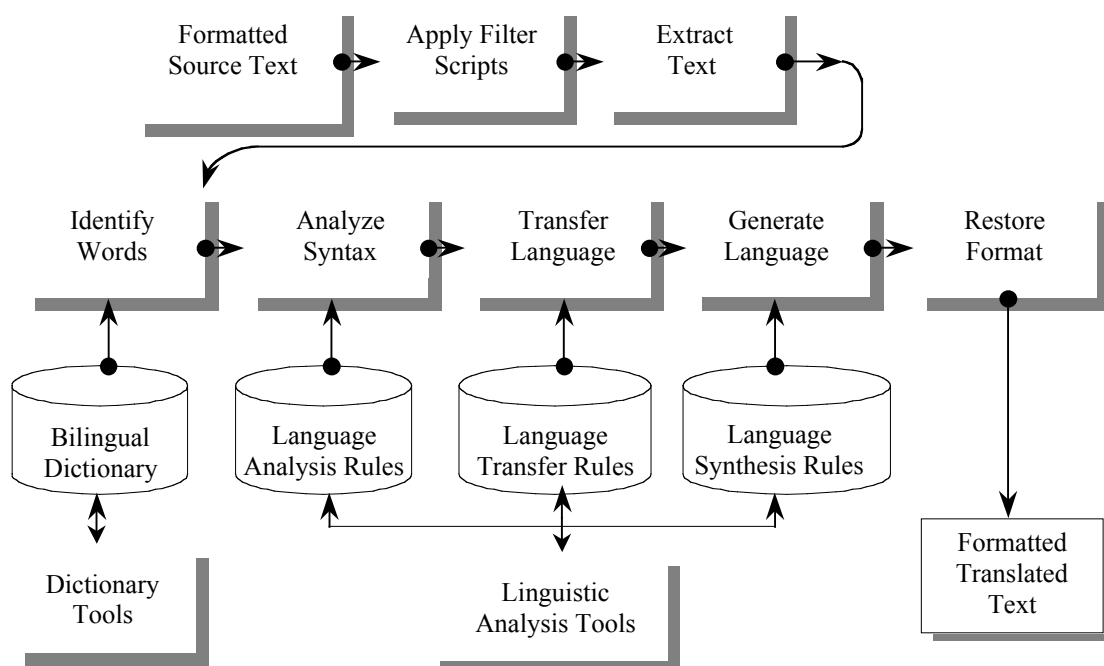


Figure 2 The TranscendRT components and their relationship [Transparent, 00]

The aforementioned architecture is implemented in the Transparent Language's Corporate and Client based products as described in the following Chapter.

3.2. SYSTRAN

SYSTRAN's (<http://www.systransoft.com/>) methodology is a sentence by sentence approach, concentrating on individual words and their dictionary data, then on the parse of the sentence unit, followed by the translation of the parsed sentence. It consists of three major groups: Dictionary, Systems Software and Linguistic Software. Each of these groups consists of a great number of modules which all work together to create a fully automatic MT system.

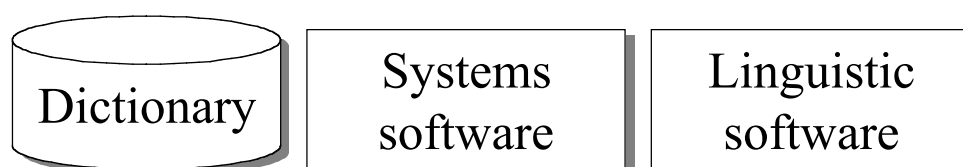


Figure 3

The Linguistic software group consists of the Parser, Target Language Translation and synthesis modules.

Parser

The most challenging aspect of any MT system is the parser, the module that analyses each sentence and attempts to build up representations of the source sentences. SYSTRAN parses with a battery of procedural modules that resolve, step by step, various syntactic and semantic relationships and assign structure within the sentence. The SYSTRAN parser is deterministic in nature, so each module makes firm decisions and passes the results on to the next module. The advantage is that every sentence, even an incomplete or malformed one, will be parsed and therefore translated. The disadvantage of such determinism, is that incorrect decisions may be passed on and compounded from module to module. SYSTRAN is able to soften this by several mechanisms that flag uncertain decisions. SYSTRAN's final step in this checking

process is a Filter program that identifies the major parse errors.

Target Language Translation Modules

After a parse of the input sentence has been constructed, algorithms for the construction of a translation are invoked. Translation information, on both the word and expression levels, is derived during dictionary lookup and the parsing phases of the translation, for use by two distinct Transfer and Synthesis modules. The Transfer component performs situation-specific restructuring, depending on the degree of difference between source and target languages. It is the only module, besides the dictionary, which relates to both source and target language, and it is rather small when the two languages are closely related.

Synthesis Module

Following this, the Synthesis module generates the target language strings which correspond to the information provided by all previous modules. Synthesis is a source-language independent module. The Synthesis modules contain sophisticated algorithms for creating specialised target language constructs, such as negation, questions, verbs with complete morphology, placement of adverbs, and articles etc.

3.3. *Lernout & Hauspie's - AppTek*

The MT software designed, developed and marketed by Applications Technology, Inc. (AppTek, <http://www.lhsl.co.uk/apptek/>) adopts the Transfer approach for parsing the source and generating the target languages. In addition to the parser, which is based on syntactic and semantic rules, AppTek's MT system includes three fully annotated dictionaries: an English lexicon, a transfer (English/Arabic or Korean or Chinese) dictionary and monolingual dictionaries for all these languages. AppTek in 1998 joined forces with Lernout & Hauspie (<http://www.lhsl.co.uk/>) and since then the technology developed by AppTek became part of a line of MT coverage that includes European language pairs.

The Machine Translation System TranSphere, has been developed as an industrial and commercial system. It was designed for large-scale volume of documentation providing for multi-sense and multi-meaning selection criteria in sentential and discourse context. Unlike small-scale off-the-shelf applications (which are semi-word-replacement in scope), the system utilises a robust parser to obtain correct meanings for words depending on their association and occurrence with each other.

The system has recently been made available for small-scale use on standalone workstations with the same robust features. The Lexicons and the linguistic rule inheritance to every lexical item helps analyse (parse), transfer and generate the output with high state-of-the-art level of accuracy. The Lexicons account for part of speech homographs as well as multi-sense, meaning and fully bilingual morphological rule assignment. The lexicons are further enhanced by Domain Specific Micro-Dictionaries (DSM) for fields such as business, finance, computer, telecommunication, aviation, military and others. The design allows for optional stacking of any or all of the dictionaries according to user specified priority. The system has a Dictionary Maintenance Utility (DMU) allowing the user to create user specific glossary and also change the system dictionary for terminology localisation. The system is also equipped with a Vocabulary Search Utility fully integrated with a morphology analyser detecting those unfound words in the system dictionary, hence allowing the user to add them using the DMU. The Windows Graphical User Interface (GUI) allows for optional synchronised viewing and an optional Translation Memory Device. The memory device retains similar phrases and sentences and later displays their post edited translations as they redundantly appear in the same text or future texts submitted to the translation process.

3.4. *LTSS*

The Lotus Translation Services for Sametime (LTSS) is developed in Lotus Ireland, a subsidiary of International Business Machines Corporation as part of the Sametime software

line. It is a Java applet that enables the translation of text, a chat session or an instant message in 17 languages.

4. Machine Translation products

There are two categories of products: for Corporate and for Private users. Both categories will be studied since the EWIN target audience ranges from consultancy firms or other companies to individual consultants or users. The Corporate (Enterprise) products offer a centralised solution to translate the documents, where all of its employees can have access. The Private user products are dependent on the user's individual platform to do the translation and the outcome is only available to that user.

4.1. L&H iTranslator Enterprise

iTranslator Enterprise is Lernout & Hauspie's – AppTek Corporate solution to document translation, for medium to large organisations. It combines machine translation technology with human translation capabilities in a way that appears seamless to the end user. There is no footprint and no application needs to be downloaded. It is a client/server system with a browser-based user interface. Documents forwarded to the “master server” for translation are routed to one or more machine translation servers, which automatically provide fast translations. For more critical documents, the master server will route translations directly to L&H's human translation services. L&H maintains a database of multilingual terminology with more than 600,000 words. The L&H iTranslator Enterprise can help you build customised dictionaries that increase the precision of translations tailored for your industry. The language pairs supported for translations are:

- English into/from French
- English into/from Italian
- English into/from German
- English into/from Spanish
- English into/from Portuguese
- German into/from French
- Italian into/from French
- Italian into/from German
- Italian into/from Spanish

4.2. L&H iTranslator Software Developer Kit (SDK)

It is a developer version that enables L&H iTranslator Enterprise to be built into any application. That can help Corporations build customised translation systems for the EWIN search results. The language pairs supported are:

- English into/from French
- English into/from Italian
- English into/from German
- English into/from Spanish
- English into/from Portuguese
- German into/from French
- Italian into/from French
- Italian into/from German
- Italian into/from Spanish

4.3. L&H iTranslator Express

iTranslator Express is a Corporate solution that enables outsourcing the translation of documents using the L&H – AppTek server farm. The users cut and paste from an email client, a word-processed document or a web site to iTranslator Express. The text is sent to the server farm, translated and sent back to the user. The solution has an important advantage that

the sophisticated translation software is not hosted at the Corporation, waving the maintenance costs.

In the user perspective iTranslator Express:

- Can translate the text after cut and paste in a text-box;
- Can translate the any Web page given its URL;
- Can translate a page automatically just by clicking a flag in the web page;
- The text can be in the ASCII or HTML format;
- Can use pre-set or custom dictionaries for more accurate translations.

The language pairs supported for translations are:

- English into/from French
- English into/from Italian
- English into/from German
- English into/from Spanish
- English into/from Portuguese
- English into/from Japanese
- German into/from French
- Italian into/from French
- Italian into German
- Italian into/from Spanish

4.4. L&H Power Translator Pro 7

L&H Power Translator Pro Version 7 is using the Barcelona translation engine to translate English to and from French, German, Spanish, Italian, Portuguese (Brazilian) and Japanese, and includes an easy to use interface and enhanced dictionaries for increased accuracy. With the use of the Web Translator feature it is easy to browse international Web sites and translate them on the fly, keeping images and links intact. It can translate documents directly within popular word processors, such as Microsoft Word or Corel WordPerfect, while preserving the document format. The L&H Power Translator Pro Editor displays the original and translated text side-by-side. Also the translation within any popular email applications, such as Microsoft Outlook, Lotus Notes, Qualcomm Eudora Pro, is feasible.

Using a sophisticated language technology, L&H Power Translator Pro translates text based on the context of the entire sentence. Accuracy is further improved by using the International ProofReader tool, which checks documents for both grammar and spelling.

It also has a look-up dictionary for more than 40 languages and over 6 million words. Using multilingual phrasebooks (for more than 36 languages) can make better translation of slang language.

In the user perspective, Power Translator includes a standalone translator module with a two-paned window: the source text goes in the top pane and the translated text appears in the bottom. The product supports HTML, RTF, and text files, and you can type or cut-and-paste text from another application into the source pane.

You can translate one sentence at a time or an entire document. An interactive mode lets you fine-tune the translation by substituting alternatives for translated words on the fly.

To improve translation accuracy further, Power Translator uses high-end linguistic rules and algorithms. In addition to general language dictionaries, the program includes over 30 industry-specific dictionaries for each language, including banking and medical. You can edit any of the dictionaries and add literal translations for specific words or phrases. These features, coupled with the fact that text is translated in the context of entire sentences, result in very effective draft translations.

A Translation Utility module integrates with Microsoft Word, Corel WordPerfect, and several e-mail packages, including Lotus Notes Mail and Microsoft Mail, Exchange, and Outlook.

This lets you create correspondence in familiar applications and translate with the press of a button.

The Web Translator module works with Internet Explorer and Netscape Navigator, translating the current Web page in your browser while maintaining formatting and hyperlinks. The user

can toggle between the original and translated versions of the Web page, and the last eight translations are saved, so they don't need to be retranslated.

L&H plans to add support for several more languages, including Chinese, Russian, Korean, and Arabic.

4.5. *L&H Simply Translating*

L&H Simply Translating can produce draft quality translations quickly and easily.

It can translate e-mails, letters, documents, and more; English to and from five of the world's most popular languages, including Spanish, German, French, Italian, and Brazilian Portuguese.

L&H Simply Translating uses sophisticated linguistic processing technology to translate the context of the entire sentence. This results in understandable, accurate, and ultimately more effective translations. It can translate from within many word processors, such as Microsoft Word and Corel WordPerfect, while preserving the original formatting.

4.6. *L&H WebTranslator*

L&H WebTranslator, quickly translates International Web pages directly within Netscape Navigator or Microsoft Internet Explorer without losing the page layout, graphics, and links. It can translate Spanish, French, German, Italian, Portuguese and Japanese web sites.

4.7. *Transparent Language - Enterprise Translation Server*

The Enterprise Translation Server is a server-based translation engine for local and remote client access to rapid machine translations. The Enterprise Translation Server enables company-wide translation of web pages, plain text and document files over both a Microsoft networking environment and a TCP/IP networking environment. The Enterprise Translation Server is ideally suited for applications, where speed and translation accuracy, are important.

With the Enterprise Translation Server, the number of users who can access translations is unlimited, and, with the Java and HTML client models included with the server, the end user's interface is accessed directly from the server.

The server receives requests for translation from clients, manages their queuing to the translation engine, and either returns the translated output directly to the end user or, in the case of a translated file, saves it to a network drive for later access by the end user. Translation requests are accepted from a wide variety of clients and translation interfaces, several included with the server, allowing flexible submission of either files or plain text for translation.

Five client models are included with the server:

1. Browser interface for web pages
2. Browser interface for text
3. Browser interface for files
4. Java applications for text and files
5. Desktop client (Windows) for files

Supporting the following formats: Rich Text Formats (.RTF), ANSI, Microsoft Word, HTML, WordPerfect, SGML, AmiPro, DOS, ASCII, PageMaker, Interleaf, ASCII, Windows Help, QuarkXpress (Tags), WordPerfect v5.1 binary and Framemaker (.MIF).

The language translations supported are:

- English into/from Spanish
- English into/from French
- English into/from German
- English into Italian
- English into Portuguese

4.8. *Transparent Language - TranscendRT SDK*

TranscendRT SDK allows custom and commercial software developers to incorporate machine translation capabilities into their new or existing applications. The TranscendRT SDK can be used to add translation capabilities to any application running on Windows 95, 98, 2000, or NT. Translation-enabled applications can run on the client desktop, network server or web server. The TranscendRT SDK contains all of the necessary DLLs and libraries to add full text or web page translation to any application.

The language translations supported are:

- English into/from Spanish
- English into/from French
- English into/from German
- English into Italian
- English into Portuguese

4.9. *Transparent Language - Easy Translator*

From text files and documents to e-mail messages and Web pages, EasyTranslator can translate any text or html-based document. It doesn't just replace individual words. It breaks down complex sentence structures, identifies parts of speech, resolves ambiguities, and synthesises the information into the components and structures of the new language. EasyTranslator's clipboard translations make feasible the translation of virtually any text. Browser-based web page translations can translate web pages complete with formatting, graphics and links.

The EasyTranslator Agent, Toolbar or Notepad can be used for Web translations. It integrates with the most popular Windows applications. The included plug-ins can make instant translations in MS Word 97, Corel WordPerfect and EudoraPro e-mail.

In the users point of view:

- Three tools integrate into the Windows desktop and help the translation:
 - Notepad View: Interacts with our work for a more accurate final output.
 - Floating Toolbar: for push-button translations.
 - Word processing plug-ins: Menu-level integration with word processing applications
 - Agent View: Access functionality from an icon in the Windows System Tray to quickly access translations
- Powerful Dynamic Dictionary Linking
- Transfer category translation (not just a word replacement program)
- Word Processing Integration for Microsoft Word, Corel and WordPerfect
- Translates Web Pages in the Browser
- Email Integration for EudoraPro 4.0

The language translations supported are:

- English into/from French
- English into/from German
- English into/from Spanish
- English into Italian
- English into Portuguese
- Japanese into English

4.10. *Transparent Language - Desktop Translator*

The Desktop Translator instantly translates documents, e-mail, web pages and speech. The user can dictate a letter, or e-mail in English and the program will translate it and even read it back in a foreign language. It has a comprehensive core dictionary and six (6) 1,000,000-word supplemental dictionaries for more accurate translations.

In the users point of view:

- Three tools integrate into the Windows desktop and help the translation:
 - Notepad View: Lets you interact with our work for a more accurate final output.
 - Floating Toolbar: for push-button translations.
 - Agent View: Access functionality from an icon in the Windows System Tray to quickly access translations
- Transfer category translation (not just a word replacement program)
- Speech to Text Dictation: Using the continuous speech technology from Dragon System, Inc. lets you enter text by speaking it.
- Voice Command & Control of the Program: Using Dragon's NaturallySpeaking technology the user can select language directions and dictionary sequences and launch translations just by speaking the proper commands.
- Text to Speech: the user can hear the source text and translations in Castilian or Mexican Spanish, French, German, Italian or US English or British English. The text to speech technology even lets the user choose between a male or female voice
- Word Processing Integration for Microsoft Word, Corel WordPerfect, and EudoraPro 4.0. The translations preserve fonts, section breaks and columns.
- Translates Web Pages in the Browser preserving its format.
- Dictionary Wizard that the user can enter key branded vocabulary like company and product names into a "do not translate" dictionary to better preserve the integrity of the source material and improve the quality of translations.
- Email Integration
- Powerful Dynamic Dictionary Linking to load industry-specific dictionaries of thousands of specialised terms and phrases common to each industry for more accurate translations

The language translations supported are:

- English into/from Spanish
- English into/from French
- English into/from German
- English into Italian
- English into Portuguese

4.11. *Transparent Language - Transcend*

Transcend enables Corporate users to get fast, accurate translations of documents, e-mail, web pages and more. Customisable dictionaries and robust linguistic technology provide the highest quality translations.

Its features include:

- Floating toolbar: for push-button translations.
- Pull-down menus. It integrates smoothly with Microsoft Word and WordPerfect by allowing you to translate documents using pull-down menus within these word processors. All page formatting and graphics are preserved in the translated document for 16 of the most popular file formats. The supported Document Formats are: Rich Text Format (.rtf), Microsoft Word, WordPerfect, Windows Help, WordPerfect v5.1 binary, DOS ASCII, Interleaf ASCII, ANSI, PageMaker, FrameMaker (.mif), QuarkXPress (Tags), AmiPro, SGML and HTMLText.
- Translate Web pages on-the-fly while retaining all graphics and formatting. Hyperlinks translate and remain functional! Supports leading Internet browsers such as Netscape Navigator and Microsoft Internet Explorer.
- Robust linguistic technology produces quality draft translations
- Interacts with Browsers such as Netscape Navigator and Internet Explorer so the user can translate on-the-fly while retaining the format, pictures and links of the original page.

- Advanced dictionary editing and dictionary sequencing technology allow you to customise and improve your translations. Teach *Transcend* to determine the correct meaning in ambiguous statements. Instruct *Transcend* not to translate trade names containing ordinary words.

There are 10 products each supporting the following language translations:

- English into Spanish
- Spanish into English
- English into/from Spanish
- English into French
- French into English
- English into/from French
- English into German
- German into English
- English into/from German
- English into Italian

4.12. SYSTRAN Personal

It can translate letters, articles, journals, e-mail and more of text not more than 5K that is cut or copied to the clipboard. The translated text can then be pasted into any application.

The supported language pairs are:

- English into/from French
- English into/from German
- English into/from Italian
- English into/from Portuguese
- English into/from Spanish

4.13. SYSTRANET

A World Wide Web browser add-on that allows the translation of web content.

The supported language pairs are:

- English into/from French
- English into/from German
- English into/from Italian
- English into/from Portuguese
- English into/from Spanish

4.14. SYSTRAN PROfessional Standard

SYSTRAN PROfessional Standard provides small businesses with powerful easy-to-use translation software that allows users to translate e-mail, letters and documents interactively or send a file off to translation with one click of a mouse.

It contains a built-in option that allows users to create a Customer Specific Dictionary with their own personalised words and expressions, that will override the product's default meanings to ensure increased accuracy in translations. It preserves the original document layout during translation and supports ASCII, ANSI, RTF, and HTML files.

The supported language pairs are:

- English into/from French
- English into/from German
- English into/from Italian
- English into/from Portuguese
- English into/from Spanish

4.15. SYSTRAN PROfessional Premium

SYSTRAN PROfessional Premium is a product for corporate users. Users can easily translate interactively on-screen or by grouping several files and sending them off to the translation queue.

SYSTRAN PROfessional Premium contains a built-in option that allows users to create multiple Customer Specific Dictionaries with their own customised words and terms, that will override the product's default meanings to ensure increased accuracy in translations.

Its features include:

- A multi-target word and phrase dictionary that contains more than 2.5 million entries.
- User-defined dictionaries with customised words and expressions that override the default built-in meanings 21 industry-specific topical glossaries for context sensitive translations
- Plug-ins for MS Word, MS Excel, MS Outlook, MS PowerPoint, Eudora E-mail and Netscape web pages
- Retains original document format
- Accepts various file formats: ASCII, ANSI, RTF, HTML, and SGML.

The supported language pairs are:

- English into/from French
- English into/from German
- English into/from Italian
- English into/from Portuguese
- English into/from Spanish
- English into/from Japanese
- Chinese into English
- Russian into English

4.16. SYSTRAN Enterprise

It is a Corporate translation tool for use in the network environment. Whether Intranet, Extranet or LAN, employees with easy access to the translation server from any world-wide client operating on the company network through an IP address. Additionally, multiple Customer Specific Dictionaries with their own customised words and terms, that will override the product's default meanings to ensure increased accuracy in translations.

Other features include:

- Multi-target word and phrase dictionaries contain more than 2.5 million entries
- Create user-defined dictionaries with customised words and expressions that override the default built-in meanings
- 21 industry-specific topical glossaries for context sensitive translations
- Plug-ins for MS Word, Eudora E-mail, Internet Explorer and Netscape web pages
- Retains original document format
- Accepts various file formats: ASCII, ANSI, RTF, HTML, and SGML.

The supported language pairs are:

- English into/from French
- English into/from German
- English into/from Italian
- English into/from Portuguese
- English into/from Spanish
- English into/from Japanese
- Chinese into English
- Russian into English

5. Machine Translation Product comparison

The machine translation products compared are for Private users and not for Corporate users. The main interest is on the languages supported, the Applications supported, the accuracy based on the product ability to support or customise the dictionaries and of-course the price. The target user is a waste specialist. The following Table is a summary of the product comparison.

Product name	Languages Supported	Applications/formats Supported	Supports Dictionaries	Price
L&H Power Translator Pro	English into/from French, German, Spanish, Italian, Portuguese (Brazilian) and Japanese	Microsoft Word, Corel, WordPerfect, Microsoft Outlook, Lotus Notes, Qualcomm Eudora Pro, Netscape Navigator or Microsoft Internet Explorer	Yes	US\$149.95
L&H Simply Translating	English into/from Spanish, German, French, Italian, and Brazilian Portuguese	Microsoft Word, Corel WordPerfect	No	
L&H WebTranslator	English into/from Spanish, French, German, Italian, Portuguese and Japanese	Netscape Navigator or Microsoft Internet Explorer	No	
Transparent Language - Easy Translator	English into/from French, German Spanish. English into Italian and Portuguese. Japanese into English	MS Word 97, Corel WordPerfect EudoraPro, Navigator and Explorer	Yes	US\$59.95
Transparent Language - Desktop Translator	English into/from Spanish, French, German. English into Italian and Portuguese	Microsoft Word, Corel WordPerfect, and EudoraPro, Navigator and Explorer	Yes	US\$129.95
Transparent Language - Transcend	Separate products for: English into Spanish, Spanish into English, English into/from Spanish, English into French, French into English, English into/from French, English into German, German into	Rich Text Format (.rtf), Microsoft Word, WordPerfect, Windows Help, WordPerfect v5.1 binary, DOS ASCII, Interleaf ASCII, ANSI, PageMaker, FrameMaker (.mif), QuarkXPress (Tags), AmiPro, SGML and HTMLText, Netscape Navigator and Internet Explorer	Yes	US\$495.00

Product name	Languages Supported	Applications/formats Supported	Supports Dictionaries	Price
	English, English into/from German, English into Italian,			
SYSTRAN Personal	English into/from French, German, Italian, Portuguese, Spanish	Clipboard	No	US\$ 69.00
SYSTRANET	English into/from French, German, Italian, Portuguese, Spanish	Netscape Navigator, Microsoft Explorer	No	US\$ 69.00
SYSTRAN PROfessional Standard	English into/from French, German, Italian, Portuguese, Spanish	ASCII, ANSI, RTF, and HTML	Yes	US\$ 350.00
SYSTRAN PROfessional Premium	English into/from French, German, Italian, Portuguese, Spanish, Japanese. Chinese into English. Russian into English	MS Word, MS Excel, MS Outlook, MS PowerPoint, Eudora E- mail and Netscape web pages	Yes	

The best tool suggested to private users for the purposes of the EWIN project is the L&H Power Translator Pro. It integrates with the most popular Web Browsers and word-processing applications. It can also support custom dictionaries to offer more accurate translations and it has a wide range of translation language pairs. Its low price makes it very competitive among the other products.

6. Web-based translation services

6.1. *FreeTranslation*

The FreeTranslation server (<http://www.freetranslation.com/>) is powered by Transparent Language's Enterprise Translation Server (<http://www.transparent.com/>). The Enterprise Translation Server uses complex technology that interprets the structure of sentences in the source language (the language you are translating from) and generates a translation based on the rules of the target language (the language you are translating to).

The highest quality translation is generated when the Enterprise Translation Server is tuned for the specific topic area of the material being translated. Using dictionaries specifically targeted to the waste industry, can substantially increase the quality of the translation.

The FreeTranslation service can translate: Spanish to English, French to English, German to English, English to Spanish, English to French, English to German, English to Italian, English to Norwegian and English to Portuguese.

6.2. *AltaVista*

The AltaVista search engine (<http://www.altavista.com/>) offers a translation service called Babel Fish (<http://babelfish.altavista.com/>) powered by the SYSTRAN translator (<http://www.systransoft.com/>). Babel Fish can:

- Translate the AltaVista French, German or Italian search results into English.
- Translate the AltaVista search results in French, German, Spanish, Portuguese, or Italian.

Babel Fish can also:

- Translate short passages to and from English to a number of languages and to and from several specific pairs of languages (for example, German to French, French to German). New language pairs are added regularly.
- Translate Web sites between specific pairs of languages. View the site in its original language side-by-side with the translation.
- Use a World keyboard to enter non-English text, where the user can use to translate or copy and paste to other applications. (For brief phrases, it's even easier to get the correct accents and characters by using Babel Fish to translate and copying the result.)

Babel Fish has imposed some limitations on the input volume, to avoid overloading the translation servers. It is designed to translate up to about 5K of text at a time (about 800 words or two standard, double-spaced pages). If you enter text that exceeds the limit or if you attempt to translate a very long Web page (over 5 K), the service translates up to the limit and then stops appending the message "Translation ends here" in the text. There is also a text passage translation limit of 1000 characters. There is also a time limit per sentence. If translation of a sentence takes longer than the allowed time, the message "Time Out" in the text is returned and the translation of the sentence is likely to be poor.

Translating languages is a very complex task. The automatic translator works best when the input text uses proper grammar. Slang, misspelled words, poorly placed punctuation, and complex or lengthy sentences can all cause a page to be translated incorrectly. Expect Babel Fish to allow you to grasp the general intent of the original, not to produce a polished translation.

Babel Fish can translate into/from English, French, German, Italian, Portuguese and Spanish.

6.3. *Systranet*

The Systranet service is available at <http://www.systranet.com/> and is using the SYSTRAN translation engine. It offers:

- Text translation using a text box
- Text translation uploading the text file
- Web page translation providing the URL

The supported languages are:

- English into/from French
- English into/from German
- English into/from Italian
- English into/from Portuguese
- English into/from Spanish
- French into/from German
- French into Italian
- French into Portuguese
- French into Dutch
- French into Spanish

With Systranet the user can select an industry specific dictionary for more accurate translations. The dictionaries supported are: Automotive, Aviation/Space, Chemistry, Colloquial, Computer/Data processing, Earth Science, Economics/Business, Electronics, Food Science, Legal, Life Science, Mathematics, Mechanical Engineering, Medicine,

Metallurgy, Military Science, Naval/Maritime, Photography/Optics, Physics/Atomic Energy and Political science.

6.4. e-lingo

The e-lingo service, which is available at <http://www.e-lingo.com/> offers a multi-lingual search function and Web surfing as well as text and email translation based on a per-transaction or monthly fee, or on a revenue-sharing model. Using the Transparent Language dictionaries, it can translate to and from English, French, German, Portuguese and Spanish. Japanese is going to arrive soon and Arabic, Chinese, Korean, and additional European languages to follow in 2001.

6.5. L&H iTranslator™ Online

It is a commerce, “pay as you go”, website using the L& H Machine Translation engine.

6.6. Globalink

The Globalink service is available at <http://www.globalink.com/> and is using the L&H translation engine.

6.7. Forthnet

Forthnet (<http://www.forthnet.gr/info/index.html>) is using the SYSTRAN translation engine to translate into/from English, German, French, Spanish, Italian and Portuguese. Will support Greek by the end of 2000.

7. Resume

The market of machine translation is very dynamic. The list therefore can't compare all available products but provides a good general overview over the market. A general recommendation for a system for eWin cannot be given since none of the products supports all European languages. The user therefore needs to select the translation system most suitable for his languages requirements.

We therefore recommend to publish this document next to a link to AltaVista's service explaining that this service is a free service with limitations and that our document compares other systems which the user might like to consider.

In parallel and with respect to the endeavours taken to commercialise eWin a link has been established to global words GmbH in 00817 Eisenach. It is planned to submit a proposal to EC for developing a translation system support Greek and to specialise the available translation system for the waste sector.

8. References

- [Transparent, 00] “Machine Translation and the Power and Purpose of “Gisting” in the Internet Era”, A White Paper from Transparent Language, <http://www.transparentlanguage.com>, January 2000
- [CE, 99] Computer Economics, Inc., June 9, 1999
- [AT, 00] http://www.access-trade.com/guides/translat_select/%20guide.shtml