Study on Evaluation of WWW MT Systems

Shinichiro MIYAZAWA (Shumei Univ.), **Shoichi YOKOYAMA** (Yamagata Univ.),
Masaki MATSUDAIRA (Oki), **Akira KUMANO** (Toshiba), **Shuji KODAMA** (Fujitsu),
Hideki KASHIOKA (ATR), **Yoshiko SHIROKIZAWA** (JST), **and Yasuo NAKAJIMA** (AAMT)

Faculty of International Cooperation, Shumei University
1-1. Daigakucyo, Yachiyo, Chiba 276, Japan
Tel. +81-47-488-2111. FAX +81-47-488-8290.
e-mail: miyazawa@cc.shumei-u.ac.jp

Abstract

Compared with off-line machine translation (MT). MT for the WWW has more evaluation factors such as translation accuracy of text, interpretation of HTML tags, consistency with various protocols and browsers, and translation speed for net surfing. Moreover, the speed of technical innovation and its practical application is fast, including the appearance of new protocols. Improvement of MT software for the WWW will enable the sharing of information from around the world and make a great deal of contribution to mankind. Despite the importance of general evaluation studies on MT software for the WWW. it appears that such studies have not yet been conducted. Since MT for the WWW will be a critical factor for future international communication, its study and evaluation is an important theme. This study aims at standardized evaluation of MT for the WWW. and suggests an evaluation method focusing on unique aspects of the WWW independent of text. This evaluation method has a wide range of aptitude without depending on specific languages. Twenty-four items specific to the WWW were actually evaluated with regard to six MT software for the WWW. This study clarified various issues which should be improved in the future regarding MT software for the WWW and issues on evaluation technology of MT on the Internet.

1. Introduction

A number of machine translation (MT) software is currently on the Japan market (Table 1). and most are for translation of the WWW (World Wide Web). This basically means that the demand for reading English homepages (HP) in Japanese is very large. The most critical issue for off-line MT is how to accurately translate text. On the other hand, WWW MT has various evaluation factors such as accuracy of text translation, consistency among various protocols and browsers, translation speed, etc. Another characteristic is that the speed of technical innovation and its practical application is rapid including the appearance of new protocols. Perfection of WWW MT software will enable the sharing of information from around the world and make a great deal of contribution to mankind. Furthermore, if MT systems in other languages are adopted in the framework of WWW MT it will be possible to translate HP into various languages which will promote mutual understanding among countries.

2 Characteristics of WWW MT

WWW MT, different from MT for regular text, should be able to perform the following:

(1) distinguish HTML tags from text
(2) understand meaning of HTML tags and translate accordingly
(3) cope with various protocols including JAVA JAVA Script, and XML
(4) cope with new WWW technologies which subsequently appear
(5) translate in a timely manner since net surfing is performed in an accelerated manner (cost control is an absolute necessity especially in the case of dial-up connection to the Internet)

3 Significance of Evaluating WWW Translation

Research on evaluating MT for text is popular. However, composite evaluation of WWW MT software in spite of its demand and importance of its development, has not seemingly been researched. Research and evaluation of WWW MT is a critical issue due to its role in international communication and its increased importance. It is strongly desired to standardize the evaluation of WWW MT.

4 Evaluation Method

Evaluation of translation quality for the text portion was not essentially conducted, since it is the same as the traditional evaluation of text MT. Items specific to the WWW instead of text were mainly evaluated. As a result, this evaluation method became applicable for general use regardless of specific languages.

5 Evaluation Items

The following items are considered as evaluation items for WWW MT software:

(1) Page setting related
- Title: Is the title for the HP translated?
(2) Font related

- Italic
- Underline: Does the translation have underline as well as in the original?
- Font type: When a different font is used in part of the original for highlighting purposes, does the relevant part in the translation use a different font?
- Font size

(3) Image related
- Letter image: Is the letter image translated?
- Clickable map: Are letters in the clickable map translated?
- Button name
- Figure reproduction
- PDF screen: Is PDF screen translated?

(4) Layout related
- Marquees: Are marquees reproduced correctly? Is the direction of marquees correct, i.e. right marquee and left marquee?

(5) Hyperlink related
- Link: Is the name of the link translated correctly?

(6) List related
- List: List should be translated as a noun phrase. Is the ending in a noun phrase?
- Symbol at the beginning of the line
- Hyphenation
- Headline size
- Consecutive number

(7) Table related
- Table: Short item name in a table should be translated as a noun phrase. Is the ending in a noun phrase?

(8) Form related
- Field type selection list: Is selection list translated?
- Drop down type selection list

(9) Frame related
- Frame: When divided in frames, is it possible to translate each frame?

(10) CGI related
- Access counter

(11) Network English related
- Internet English: Are expressions and terms specific to the Internet translated?
- E-mail English: Are expressions and terms specific to E-mail translated?

Eleven HPs with the above characters unique to the WWW were searched on the Internet and used as test pages (Table 2: hereafter indicated as HP1-HP11 for reference to each page.) Main items should be covered here, although other evaluation items may be considered.

6 Translation Results and Evaluation

Six MT software was used to translate English into Japanese. Analysis and evaluation of translation results were as follows:

6.1 Page setting related

(1) Title

HP titles remained in English in the translation results for three MT software. The other three software was able to translate titles. Titles should be translated because it is possible to read titles from HTML tags.
One software which was able to translate, however, only translated partially as follows:

・Original 1: Terms and Definitions (HP11)
・Translation 1-1: 項と Definituons

This might have occurred due to capitalization, but it should be translated properly even with capitalization because it is a title.

6.2 Font related

(1) Italic

In the translation results shown in Figure 6 (HP8), two software did not reproduce italics. All other software properly showed italics.

(2) Underline

The two software in Figure 6 (HP8) did not reproduce underline. Among other software, some underlines were missing in one software as follows: (「強調された」should have been underlined too.) It may be an unnecessary addition but <> did not correspond properly in the translation results, although <U> and </U> were not actual tags.

・Original 2: ・Use < U > for underlined text < /U > (HP8)
・Translation 2-1: ・使用 U> は> 強調された テキスト/U のために

Underlined parts should have been understood. Since "<>" is mostly used as a pair, the correspondence should be understood and utilized positively to carry out syntax analysis for translation.

(3) Font type

For places with different English fonts such as Original 3, none of the software used a variation of fonts in the translation results. Since different fonts are used for necessity, it is better to use different fonts in the translation. When there is no font available similar in appearance as the English font, it will be effective to change the text to some other generally-used font of that language.

・Original 3: **<FONT FACE="comic sans ms" > for "Comic Sans MS" text < /FONT >** (HP8)

(4) Font size

Three software changed font sizes in the translation results, when font sizes were changed in the original. Font sizes should be changed properly since changes in letter size are distinguished in tags.

・Original 4: n can be any number between 1 and 7. Normal font size is 3. This font size is 1 (HP8)
・Translation 4-1:n は、1 と 7 の間にいくらかの(どんな)ナンバーであることができます。普通のフォント・サイズは 3 です。このフォント・サイズは 1 です。

(5) Proposal to overall font related matters

Font related information should be reflected in translation by all means. It may be necessary to improve the translation engine of MT software because one English word may be separated into two or more in translation.

6.3 Image related

(1) Letter image

None of the software translated the letter parts regarding GIF files in Figure 1 (HP9). Since they are simple letter images, they should be easily recognized using the

OCR function.

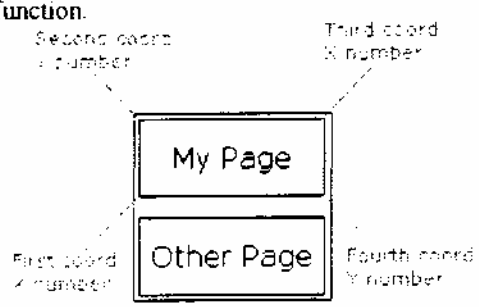

Figure 1: Letter Image (extracted from HP9)

(2) Clickable map
None of the software translated clickable maps in Figure 4 (HP1).

(3) Button name
Button names remained in English for all software. However, button names were described in HTML source. For example, buttons like "Go!" and "Search" in Figure 4 (HP1) were described as follows in the source:
<input type="submit" value="Go!">
<INPUT TYPE=SUBMIT VALUE="Search">
Button names can be translated if this "VALUE=" part is understood.

(4) Figure reproduction
All software reproduced figures and images in the original HP.

(5) PDF screen
The letter part in PDF in HP6 remained in English for all software. Since there is text in PDF for the letter part, however, it can be translated if MT software can cope with PDF protocol. This should be made possible in the future since PDF protocol is becoming more widely used.

(6) Proposal on overall image-related matters
Letters in image-format are frequently used in HP, therefore many parts cannot be translated without translating the letter image. Some buttons are only with image screen without "VALUE=". Using current technology, it is difficult to translate clickable maps, but due to their frequent use, some effort is necessary. Translation of letters in image-format is a fundamental problem in WWW MT.

One option is to recognize letters in image-format by incorporating the OCR function. Although OCR's recognition level is not perfect, translation will be possible to some extent through interaction with users, i.e. confirm with users if there is no misspelling in the translation results.

Another option is for HP creators to write comments in text equivalent to letters in image-format. This may be a burden to HP creators. If they understand that their HPs will be read worldwide by doing this, it may give some incentive to them. It may be necessary to promote this as netiquette for HP creators. Or in order not to put the burden on HP creators, it may be necessary to research and develop WWW protocol to describe such comments automatically.

Even if it becomes possible to translate the text in figures, a problem may occur including location or placement of text in figures. This issue needs to be reviewed. Considering the case to take (refer to) shop names or street names during travel, both the original and the translated results may be indicated in many cases.

6.4 Hyper link related

(1) Link name
Individual link names in Figure 4 (HP1) were not translated which resulted in the following:
1. Some links were not translated.
2. The link order changed.
3. One link became two.
4. Two links, on the other hand, became one.
5. Multiple links became one sentence.

・Original 5: Yellow Pages Maps Farefinder Reservations by Preview Travel Related Books (HP1)
・Translation 5-1: Preview Travel 関連づけられた本による 職業別広告欄 Maps Farefinder Reservations

・Original 6: Help Add URL Advertise on Excite Press Releases Jobs@Excite (HP1)
・Translation 6-1: Excite Press Releases Jobs@Excite の助け (cf. 4) Add URL Advertise
・Translation 6-2: URL が Excite プレスリリース Jobs@Excite で広告を出すのを言い足すのを助けなさい

In the above translation, the succession of phrases between <A> and </A> was interpreted as one sentence by mistake. Translation between the tags should be made possible by reading tags for link (<A HREF> and </A>) well. It will be necessary to improve MT software for this purpose. A problem with HTML is that HTML tags are not able to describe logically that "Yellow Pages", "Maps", "Farefinder", etc. are independent items. As a result, it cannot distinguish from the instance where it is actually one sentence.

It was not researched in this study, but regarding an instance where there is a link name in one sentence, further study will be necessary to review if translation should be as a link name or based on its role in the sentence. For example, there is an instance where a link name is located in the verb area, it works as a verb in the original, and it is actually a link. The way it should be expressed as a translation needs to be reviewed.

6.5 List related

(1) List
Regarding the translation results of the list of "Related Topics" on the left side (Figure 5) of HP3, some software was not able to translate even half of the list, because the words with capitalized first letters were not translated. It seems that words with initial caps were regarded as proper nouns, and as a result, were not translated. Since these words are frequently used in lists, MT software's effectiveness is not utilized enough when half of the words are not translated. These words should be translated; if users doubt the translation, they will just have to review the original.

The following example indicates failure by translating proper nouns. Words may not be translated when all letters are capitalized.

・Original 7: UH Admissions (UH is an abbreviation of University of Hawaii.) (HP3)
・Translation 7-1: あの、承認 ("あの" is a demonstrative noun in Japanese.)

The next example shows the case where a list is not in noun phrases. Lists should generally be in noun phrases.

・Original 8: InterNIC Guide to U.S. Universities (HP3)
・Translation 8-1: 「米国大学に対する InterNIC 導く」

(2) Symbols in the beginning of lines
The symbols in the beginning of lines "・" were omitted in five different software.

・Original 9: ・Abbreviations Used at the UH (HP3)
This means that MT software is not able to deal with HTML tags (<UL><LI>.....</UL>).

(3) Headings in noun phrases
Headings are usually in noun phrases. The next example shows the instance where a heading was not translated into a noun phrase. Headlines between symbols in the beginning of lines ("・" and "-") which indicate items should especially be translated into noun phrases.

・Original 10: ・ping - (HP10)
・Translation 10-1: ・確認しなさい. -

(4) Hyphenation
One software omitted one hyphen in the heading. All other hyphens in HP10 remained in the translation by this software.

Original 11:
・flame - An electronic message that is particularly hostile. Can also be a verb: "Whooeee! I posted a rude cat joke to my company's cat-lovers mailing list, and wow, did I get flamed!" (HP10)

Translation 11-1:
・燃え立ってください。特に対立的な電子通信。さらに動詞でありうる:「Whooeee! 私は、会社の猫好きメーリング・リストへの無作法な猫冗談、およびワウを記入しました、行った、私は得ます、燃え立った!」

The following shows an example where "YMMV - " in the heading moved and was not located at the front in the translation results. Hyphenation also disappeared.

Original 12:
・YMMV - Your Mileage May Vary (taken from a disclaimer that legally must be given any time automotive fuel efficiency ratings are used in U.S. advertisements) (HP10)

Translation 12-1:
・あなたの走行マイル数 5 月、(合法的に自動車の燃料効率評価が合衆国広告で使われるどんな時でも与えられなくてはならない断り書きから取り出される) Vary を YMMV してください

In the translation results of HP10, three software omitted spaces before and after hyphens although the original had them. It is important to have spaces before and after hyphens, otherwise hyphens may easily be misunderstood as macrons in Japanese.

(5) Size of headings
All software used a larger sized headings in translation when a larger size was used in the original as shown in Figure 6 (HP8). It is important in terms of ease of viewing to use a larger size for headings. Especially in the case of WWW MT software, headings are often used to demonstrate the general contents of the HP at a glance, therefore ease of viewing in translation is particularly important.

(6) Consecutive numbers
HP7 uses HTML's function of consecutive numbers. Five software was able to reproduce them. Two of them changed periods to punctuation marks as shown in the following example. It is desirable to use periods instead in the case of consecutive numbers.

・Original 13: 1. (HP7)
・Translation 13-1: 1。

One software did not reproduce consecutive numbers at all, and they were omitted. Consecutive numbers need to be reproduced in translation because they have an important role in terms of ease of viewing.

6.6 Table related

(1) Tables
Sentences in tables are usually in noun phrases. Most parts in the tables of Figure 3 (HP2) were translated into noun phrases, but some parts were not. The following examples indicate instances where item names in tables were not in noun phrases, or periods were used at the end of sentences.

・Original 14: Applied Language Studies: Combined Studies (MA) (HP2)
・Translation 14-1: 実用的な語学:研究(MA)を結び合わせました

・Original 15: Taught (HP2)
・Translation 15-1: 教えられた.

・Original 16: English: Modern Literature and Culture (MA) (HP2)
・Translation 16-1: イギリスである:現代の文学と文化(MA)

6.7 Form related

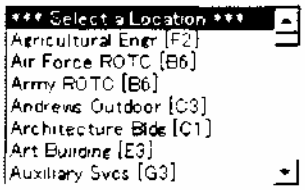

Figure 2: Field type selection list (extracted from HP4)

(1) Field type selection list
This shows the translation results of field type selection lists such as in Figure 2 (HP4).
One software did not translate and leave them as they are in English, and others translated them. However, asterisks were not used before and after the title of selection list for two software. (There was an instance of inserting a comma in the middle.) It is desirable to reproduce asterisks correctly because they are frequently used as shown in the original, especially in headings of selection lists.

・Original 17: *** Select a Location *** (HP4)
・Translation 17-1: ***、位置***を選択する。

Regarding two software, only the following item was not in a noun phrase although others were. Items in the same selection list should be either in noun phrases or not.

・Original 18: Art Building [E3] (HP4)
・Translation 18-1: Art は[E3]構築しています

(2) Drop down type selection list
The drop down type selection list in HP1 was translated by five software, and remained in English for one software.

6.8 Other problems specific to the WWW

(1) Marquees (layout related)
Only one software did not reproduce marquees of HP7 on the screen since the tag part was not properly processed. Translation was carried out by all the software.

(2) Frame
The frame in HP4 was reproduced by four software, however for two of them, each frame was printed on a separate page. One of the two did not reproduce maps.

(3) Access counter (CGI related)
The access counter in HP5 was reproduced by all software. However regarding the access counter in HP11, four software indicated zero regardless of the number on the original HP. One software indicated one in the access counter.

The original page shows five-digit numbers, not zero or one. The HTML source of this page indicates:
You are visitor number < IMG SRC="/cgi-bin/counter/555babe">
which means that the CGI program called 555babe displays the count number in GIF files. The relevant part in HTML file of the translation indicates:
あなたは訪問者の数である< IMG SRC="http://www.geocities.com/cgi-bin/counter/555babe">

which means that exactly the same CGI program must be working. Programs on the server side may process differently for regular and non-regular access. It is not possible to judge more than that as long as the structure of the 555babe program is not clarified.

In one software, the following original sentence which was located by the side of the access counter was not translated and remained in English.

・Original 19: You are visitor number (HP11)

(4) Software Crash
One software stopped and crashed during translation of HP10.

6.9 Network English

These analyses are referring to general problems which are not limited to specific languages to be translated.
The theme of "how to correctly translate Internet English" dealt with here is a problem depending on Japanese. Although it is a problem depending on the specific language, English words specific to the Internet should especially be translated correctly because we are dealing with WWW MT. The translation results, however, varied, i.e. some cases were translated correctly and others remained in English.

(1) Internet English
The following shows HP7 and HP11 which are examples with headings failing to be translated. For each HP, one software did not translate correctly, while others did.

The following examples show instances translated into different meanings from those used in the Internet.
For example, basic Internet terms such as "web page" and "Internet Explorer" should be listed as standard terms in the dictionary of MT software. WWW MT software should be given priority in translating words related to the Internet and used frequently as Internet terms.

・Original 20: Web Page (HP11)
・Translation 20-1: 力者ページ

・Original 21: Hypermedia (HP11)
・Translation 21-1: 過メディア

・Original 22: Using Window 95 and Microsoft Internet Explorer (HP7)
・Translation 22-1: 使う Windows 95 および「マイクロソフト・インターネット探検家」

・Original 23: Spam (or Spamming) (HP11)
・Translation 23-1: 何度も送(そうでなければ「何度も送る」)

(2) E-mail English
HP10 explains E-mail terms. "BTW" shown below is an abbreviation specific to E-mail, and only one software was able to translate this term. "FYI" is also an abbreviation and two software succeeded in translation. Since "BTW" and "FYI" are headings in this case, it is correct as a sentence if these words remain in English. However, in terms of translation ability, it was demonstrated that these software was able to translate them not only when they were used in headings but also in the middle of sentences.

・Original 24: BTW - By The Way (HP10)
・Translation 24-1: ところで-ところで

・Original 25: FYI - For Your Informaion (Product B) (HP10)
・Translation 25-1: ・参考のために-ご参考までに

The next example is an instance of failed translation

This Japanese translation has the wrong pronunciation.

・Original 26: ping - (HP10)
・Translation 26-1: ピューン-

6.10 Translation quality related

Quality of translated sentences is a critical problem for general MT as well, and not a problem specific to the WWW, however the following indicates some points noted.

(1) Success
・Original 27: University of Hawaii at Manoa (HP4)
・Translation 27-1: ハワイ大学マノア校

"Manoa," a proper noun, was properly translated in Katakana used to express a foreign sound in Japanese.

(2) Failure
・Original 28: Office Phone No. (HP3)
・Translation 28-1: オフィス電話…いいえ

No. was translated to mean denial. Since "Office Phone" precedes it, it should be translated as "number."

・Original 29: Add BGCOLOR="colour name or code" to the MARQUEE tag (HP7)
・Translation 29-1: または…バックグラウンドカラー=」 色の名を加えなさい…コード」…

" " is expressed in 「 」 in Japanese, but its correspondence was not handled properly. Since 「 」 is a pair, its correspondence should be given priority.

・Original 30: ・Use < U > for underlined text < /U > (HP8)
・Translation 30-1: ・ 使用 <U> 強調されたテキストのために </U>

"underline" was translated as "emphasis (強調)" because of too-much considering the meaning of "underline."

7 Preparation of Test Page for Standard Evaluation

It required substantial time to look for HPs suitable for evaluation. Even if found, the contents change or it may be discontinued in time; therefore it gets difficult to conduct tests in an uniformed and general manner. Since this evaluation clarified issues of WWW MT, a general test site for standard evaluation is under construction. HP for this site is in English. When testing of WWW MT software is desired, the quality will be tested by accessing this site and running MT.

8 Conclusion

Table 3 indicates evaluation results for each product. Issues to be improved in the future for WWW MT software are becoming apparent. Regarding the items that none of the software can properly translate, higher performance should be developed after this. However, tasks at least one software was able to perform should be able to be adopted by other software manufacturers.

Based on these factors, the following themes will be pursued as issues for the evaluation of Internet MT:

1) Enrich the test site for standard evaluation as described above.

2) Improve quality of WWW MT.
(1) Propose to manufacturers regarding the improvement of MT software.
(2) Propose to users (page creators) regarding ideas on HTML description.
(3) Propose alternative ideas regarding matters difficult to improve under current technology.
Propose to add the relevant text information in HTML as a comment at the same time, since letters in image format such as letter images and letters in clickable maps are very difficult to translate.
(4) Measure frequency (importance) of WWW function usage.
Provide resources to prioritize for effective improvement.
(5) Prove uniformity of terms in translation in the same HP.
(6) Propose a WWW protocol easy to translate.
When human beings read web pages, apparent emphasis (i.e. *** *** in Original 17) or apparent structure (i.e. succession of phases in Original 5 or Original 6, hyphenation in Original 11 or Original 12) is adequate. However for processing by machine using MT, it may cause a misunderstanding or misinterpretation since meanings are not defined logically. It will be necessary after this to write web pages on the condition of machine processing like MT. For this purpose, such vague expressions only understood by human being should be newly specified as tags.

3) Evaluation of translation speed
Translation speed is especially important for online WWW translation.

4) Browser
Investigate browsers for WWW MT software.

5) Evaluation of Internet MT except WWW
Evaluate translation quality of E-mails, telnet screens, chat screens, etc. in particular.

With on-going internationalization and development of information technology including the computer network and international groupware, demand and expectations on the Internet MT system will grow. Since software development will be promoted through evaluation, development of evaluation technology is very important. We hope this study will be of help for that purpose.

References

Fujita, K., Kawagoe, M., and Kurisaki, M. (1997). "English-Japanese Translation Software for Internet Browsers". National Technical Report, Vol. 43, No. 6, pp. 90 - 98 (in Japanese).

Murata, T., Yamamoto, H., and Nagata, J. (1996). "Realtime Machine Translation System for World Wide Web". IEICE Trans., Vol. J79-B-I, No. 5, pp. 364 - 372 (in Japanese).

W3C (1998). "HTML Specification". http://www.w3.org/TR/REC-html40/

MT Software for Windows (English → Japanese)

Products(Japanese)	Ver.	3.1	95	98	NT	Company	Price(¥)
EtoJ INTERNET PLUS	4.0		○	○	○	Catena Co.	12,800
Translation Surfin+(翻訳サーフィン+)	5.0		○	○	○	Fujitsu Ltd., Fujitsu Middleware Ltd.	12,800
Translation Surfin (翻訳サーフィン)	5.0		○	○	○	Fujitsu Ltd., Fujitsu Middleware Ltd.	8,800
The Translation Internet (The 翻訳インターネット)(English → Japanese)	3.0		○	○	○	Toshiba Co.	16,800
The Translation Internet (The 翻訳インターネット)(English → Japanese)	3.0		○	○	○	Toshiba Co.	12,800
SurfMate	1.0		○			Toshiba Co.	8,400
The Translation Basic (The 翻訳ベーシック)	1.0		○	○	○	Gakken Co., Three-A Systems Co., LTD.	3,980
Internet the Kokusaijin (インターネット ザ・国際人)	4.1		○			Sanyo Information Business Co.	9,800
NetSurfer/ej	3.0	○	○			Nova Co. Ltd., Yamatane Co., NEC Information Service, Ltd.	24,800
Transpad	4		○	○	○	AI Logic	14,800
Transpad NN	2		○	○	○	AI Logic	6,800
PENSEE for Internet	2.0		○		○	Oki Software Co., Ltd.	9,800
WorldNet (English ⇔ Japanese)	4.0		○	○	○	Kodensha Co. Ltd.	12,800
King of Internet Translation (インターネット翻訳の王様) (English → Japanese)	3.0		○	○	○	IBM Japan	7,800
Dr. SURF	2.0		○			Media Vision	12,800

MT Software for Macintosh (English → Japanese)

Products(Japanese)	Ver.	Company	Price(¥)
EtoJ INTERNET PLUS	4.0	Catena Co.	12,800
NetSurfer/ej	3.0	Nova Co. Ltd., Yamatane Co., NEC Information Service, Ltd.	24,800
Transpad	2	AI Logic	14,800
Dr. SURF	2.0	Media Vision	12,800

Table 1: List of Internet MT Software (Japan) (extracted from http://www.bekkoame.ne.jp/~oto3/jhonyaku.html)

HP	URL
HP1	http://city.net/regions/north_america
HP2	http://www.ukc.ac.uk/prospectus/postgraduate/pgprogs/alphalist.html
HP3	http://www.hawaii.edu/cgi-bin/webph?DB=csnet-ns
HP4	http://www.hawaii.edu/campusmap/
HP5	http://www.uhh.hawaii.edu/
HP6	http://www.hawaii.edu/campusmap/uhmmap.pdf
HP7	http://subnet.virtual-pc.com/li542871/marquee.htm
HP8	http://subnet.virtual-pc.com/li542871/text.htm
HP9	http://www.geocities.com/SiliconValley/Heights/1288/imaps.html
HP10	http://www.webfoot.com/advice/email.jargon.html?Email
HP11	http://www.geocities.com/FashionAvenue/4869/desc.html

Table 2: Tested Homepages

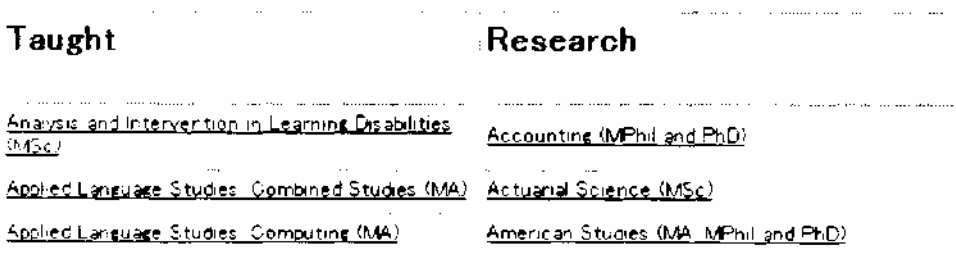

Figure 3: Table (extracted from HP2)

Table 3: Evaluation of WWW Machine Translation

	Product	A	B	C	D	E	F	Evaluation of translation				Note
Relevant HP	Evaluation item							○	△	×	Total	
HP1	Clickable map	×	×	×	×	×	×	0	0	6	6	
	Title	×	×	×	○	×	×	1	0	5	6	
	Figure reproduction	○	○	○	○	○	○	6	0	0	6	
	Drop down type selection list	×	○	△	○	○	○	4	1	1	6	
	Link name	△	△	×	△	×	×	0	3	3	6	
	Button name	×	×	×	×	×	×	0	0	6	6	
HP2	Title	○	×	○	○	×	×	3	0	3	6	
	Table (in noun phrases)	△	△	△	△	△	△	0	6	0	6	
HP3	Title	○	×	△	○	×	×	2	1	3	6	
	Drop down type selection list	×	○	○	○	○	○	5	0	1	6	
	List	○	○	○	○	△	×	4	1	1	6	
	Button name	×	×	×	×	×	×	0	0	6	6	
HP4	Adaptability to frames	○	○	○	○	○	○	6	0	0	6	
	Print frames in the same page	×	×	○	×	×	×	1	0	5	6	
	Field type selection list	×	○	○	△	○	△	3	2	1	6	
	Button name	×	×	×	×	×	×	0	0	6	6	
HP5	Reproduction of access counters	○	○	○	○	○	○	6	0	0	6	
HP6	PDF screen	×	×	×	×	×	×	0	0	6	6	
HP7	Title	△	×	○	○	×	×	2	1	3	6	
	Reproduction of symbols	○	○	○	○	○	○	6	0	0	6	
	Marquee	○	○	○	○	○	○	6	0	0	6	
	Actual marquee (left and right)	○	×	○	○	○	○	5	0	1	6	
	Consecutive numbers	×	○	○	○	△	△	3	2	1	6	This is not a problem specific to WWW translation
HP8	Under line	○	×	○	×	○	○	4	0	2	6	
	Title	○	×	○	○	×	×	3	0	3	6	
	Italic	○	×	○	×	○	○	4	0	2	6	
	Font type	×	×	×	×	×	×	0	0	6	6	
	Heading size	○	○	○	○	○	○	6	0	0	6	
	Letter size in sentence	×	○	○	×	×	○	3	0	3	6	
HP9	GIF file	×	×	×	×	×	×	0	0	6	6	
HP10	Title	○	×	○	○	Software crash	×	3	0	3	6	
	E-mail English	△	△	△	×	Software crash	○	1	3	2	6	
	Headings in noun phrases	△	△	△	△	Software crash	○	1	4	1	6	This is not a problem specific to WWW translation
	hyphens	○	△	○	○	Software crash	○	4	1	1	6	This is not a problem specific to WWW translation
	Spaces before and after hyphens	×	×	△	○	Software crash	×	1	1	4	6	This is not a problem specific to WWW translation
HP11	Access counters	×	△	△	△	△	△	0	5	1	6	
	Internet English	△	○	○	○	△	○	4	0	2	6	
	Headings in noun phrases	△	○	○	△	○	△	3	3	0	6	
	Title	△	×	○	○	×	×	2	1	3	6	

○: success △: partial success ×: failure

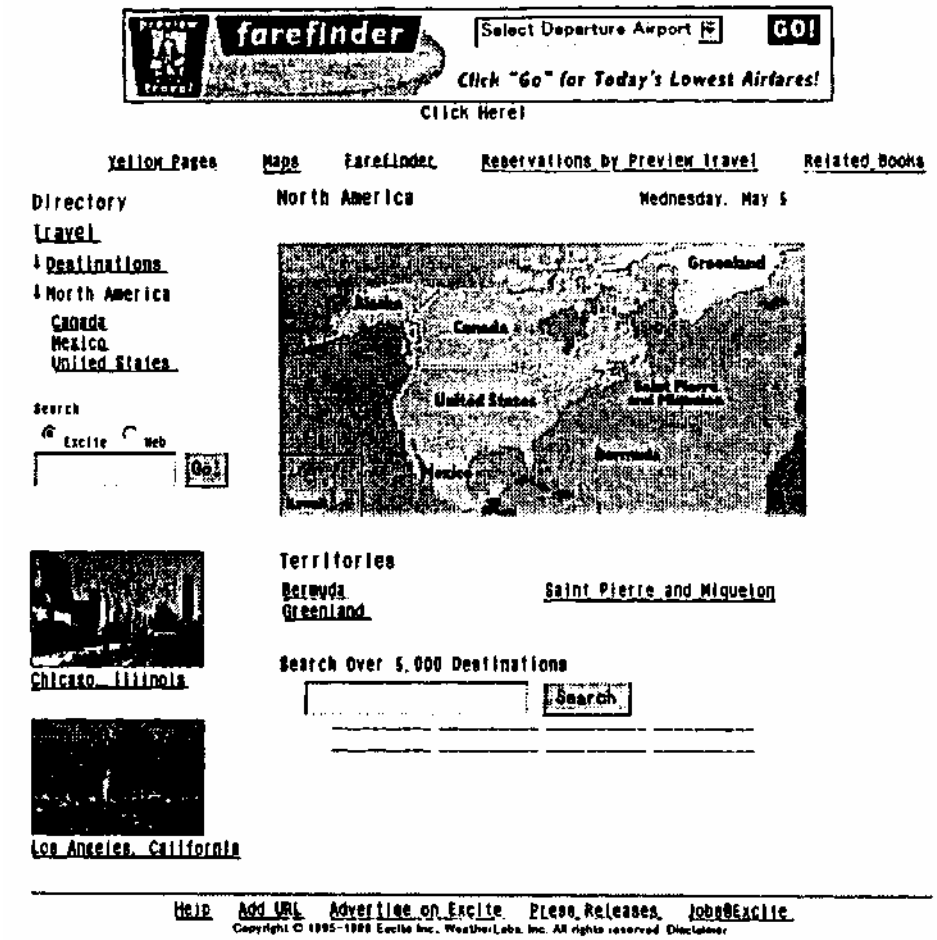

Figure 4: HP1 (extracted from http://citi.net/regions/north_america)

3. Changing the type and size of text

- Use < B > for bold text< /B >
- Use < I > *for italic text* < /I >
- Use < U > for underlined text < /U >
- Use < TT > for evenly spaced typewriter text< /TT >
- Use < FONT SIZE=n > to get different text sizes < /FONT >

n can be any number between 1 and 7. Normal font size is 3

4. Changing the text font

This only works for people viewing your pages with Internet Explorer or Netscape Navigator 3

You can change the text font by typing

<FONT FACE="times new roman" > for "Times New Roman" text < /FONT >

<FONT FACE="courier" > for "Courier" text < /FONT >

<FONT FACE="comic sans ms" > for "Comic Sans MS" text < /FONT >

Figure 5: List (extracted from HP3)

Figure 6: Font (extracted from HP8)