

A Survey of Computer Usage in Japanese Translation

James BREEN
Monash University
Clayton 3800, Australia
jimbreen@gmail.com
June 2007

Abstract

This report contains a summary of the results of a survey of 171 translators working with the Japanese language. The survey was carried out in May 2007, and investigated aspects of the usage of computer and communications technology in their work.

1. Introduction

In preparation for a presentation on Japanese Translation and the Computer to the International Japanese English Translation (IJET) Conference in Bath (UK) in June 2007, the author carried out a survey of translators working with Japanese to determine the usage of computers, communications and online services in their work, the changes in that usage over time, and their opinions on various aspects of the interaction of computers and related technology with their work.

The survey was carried out over a three-week period in May 2007 using a WWW-based system developed by the author. (The questionnaire can be seen at: <http://www.csse.monash.edu.au/~jwb/cgi-bin/enquete.>) Invitations were issued on the *Honyaku* mailing list, which is a major discussion forum for people working in Japanese translation, with the request that the invitation be passed on to other related mailing lists within national translation associations, such as JAT (Japan), ATA/JLD (USA) and ITI/J-NET (UK).

171 usable responses were received to the questionnaire. This report summarizes the information provided in those responses.

2. Profile of Respondents

The initial questions were to determine a broad profile of the respondents; their nationality, country of residence, length of time translating, etc. The results are summarized in the following tables.

Country	USA	Britain	Japan	Australia	New Zealand	Canada	Germany	Other
Nationality	84	28	20	14	5	5	3	7
Country of Residence	49	20	74	9	3	2	1	9

Table 1: Nationality and Country of Residence

Time translating	0-2 years	3-5 years	6-10 years	> 10 years
(%)	13.6	13.1	19.3	54.0

Table 2a: Period earning income as a translator

Working Mode	In-house	Freelance	Mix of in-house/freelance
(%)	17.6	72.2	10.2

Table 2b: Working Mode: in-house/freelance

Translation Direction	from Japanese	into Japanese	a mix
(%)	75.6	4.5	19.9

Table 2c: Translation Direction

Of the 36 respondents who reported translating in both directions, 13 were Japanese nationals for whom the average proportion of translation into English was 46%. The other 23 reported translating into English an average of 81% of the time.

Translation as an Occupation	Sole Occupation	Main Occupation	Second Occupation
(%)	65.3	22.7	11.9

Table 3: Translation as an Occupation

A wide range of other occupations were reported, among which teaching (Japanese and other languages), editing/proof-reading, and consulting were the most common.

3. Computer Usage and Expenditure

Respondents were asked for the Operating System they used, the age of the computer they while translating, and the average expenditure on computer-related facilities in recent years.

Operating System (Primary Computer)	Windows	Macintosh	Linux/Unix
(%)	80.7	19.3	0.0

Table 4a: Primary Operating System

Operating System (Secondary Computer)	None	Windows	Macintosh	Linux/Unix
(%)	78.9	10.5	8.8	1.8

Table 4b: Secondary Operating System

Computer-related Expenditure	Hardware	Software	Electronic Dictionaries	Internet Connection	Other
Number of Respondents	147	139	80	150	78
Average Annual Expenditure (\$US)	1967	844	393	730	501

Table 5: Computer-related Expenditure

4. Document Delivery from Clients

Respondents were asked for the approximate percentages of documents they received in various forms and media, and for comments on any changes in the patterns of format and media in recent years.

Delivery method media	Hard-copy - mail	Hard-copy - facsimile	Electronic copy - email	Electronic copy - disk/CDROM/etc.	Other
Respondents	83	73	166	45	15
All respondents (%)	7.6	3.5	82.5	4.7	1.6
Respondents for media type (%)	15.8	8.5	85.7	17.8	18.6

Table 6a: Document-delivery method and media

Electronic Documents - file types	Word-processing	PDF	Spreadsheet	Database	Presentation	WWW page	Plain text	Other
Respondents	166	140	130	2	113	39	26	22
All Respondents (%)	49.8	29.6	10.1	0.2	5.7	2.0	0.9	1.7
Respondents for file type (%)	58.8	41.5	15.2	15.0	9.8	10.2	6.5	15.3

Table 6b: Electronic Document file type

Comments from respondents included:

- remarks on the significant decline in the use of hard-copy. Several said they refused work unless it was in electronic form.
- there is an increase in downloading of documents, e.g. from clients' document management systems, the Patent Office, etc.
- many commented on the increase in the use of PDF files, although a few commented that they saw PDFs declining in favour of Word documents
- several comments that requests to translate Powerpoint documents were increasing.

5. Dictionaries and Glossaries

Respondents were asked about their use of dictionaries and glossaries, including usage patterns of paper dictionaries, electronic dictionaries and online dictionaries.

Paper Dictionary usage	Rarely use	Use but less than before	Use just as much as before	Use is increasing	Only use
(%)	60.9	30.2	7.1	0.6	1.2

Table 7: Paper dictionary usage

Paper dictionaries mentioned by respondents included: Kenkyusha JE - 4th and 5th editions (45 mentions), Nelson (28), Daijirin (9), Kojien (6), The Kanji Dictionary (5). In addition some domain-specific dictionaries were mentioned several times, such as Blacks (law) and Stedman (medicine).

Electronic dictionary usage	Rarely use	Use but less than before	Use just as much as before	Use is increasing	Only use
(%)	34.3	16.6	10.7	10.1	28.4

Table 8: Electronic dictionary usage

The local-file-based electronic dictionaries which were mentioned most often included the Kenkyusha JE 5th edition (24 mentions), Eijiro (17), EDICT (16), Kôjien (11) and Daijirin (5). The Jamming package was mentioned quite often as a client for searching dictionaries. Most popular among the hand-held dictionaries were models by Seiko (13 mentions), Canon (11) and Sharp (3).

Online dictionary usage	Rarely use	Use but less than before	Use just as much as before	Use is increasing
(%)	14.8	4.7	24.3	55.2

Table 9: Online dictionary usage

The dictionary server which was mentioned most often was Eijiro/ALC with 81 mentions, i.e. almost half the respondents. The other sites which were often mentioned were WWWJDIC (41), Glova (34), Kenkyusha (16), Goo (8), Jeffrey's (8) and Yahoo (6). A wide range of subject-specific online glossaries were mentioned, although no single glossary dominated.

Finally, respondents were asked to assess their overall use of dictionaries (as opposed to other sources of information).

Overall dictionary usage	Declining	Staying much the same	Increasing
(%)	35.5	52.1	12.4

Table 10: Overall dictionary usage

6. WWW/Archive/Corpora Searching

A series of questions were asked to determine the use of WWW searches, mailing-list archives, etc. as an aide to translation.

Use of WWW search engines	Never	Occasionally	Often	Very regularly	Many times a day
(%)	1.8	2.3	9.9	22.8	63.2

Table 11a: Use of WWW Search Engines

Usefulness of WWW in translation	No use	Limited	Quite useful	Very useful	Indispensable
(%)	2.3	0.0	10.5	17.0	70.2

Table 11b: Usefulness of WWW Search Engines

WWW usage over time	Decreasing	Same level	Increasing
(%)	3.5	35.7	60.8

Table 11c: Trend in WWW Search Engine Usage

Mailing list archive usage	Never	Occasionally	Often	Very regularly	Many times a day
(%)	26.3	56.7	10.5	4.7	1.8

Table 12a: Use of Mailing List Archives

Mailing list archive usefulness	No use	Limited	Quite useful	Very useful	Indispensable
(%)	18.1	38.0	26.3	11.7	5.8

Table 12b: Usefulness of Mailing List Archives

Respondents were asked if they searched any other corpora. Less than a third nominated such corpora, with searches of prior translations being the most commonly mentioned, followed by Wikipedia.

7. Translation Memory Systems

A series of questions was targeted at the use of Translation Memory Systems, including patterns of usage, reasons for use, productivity gains, and issues of exchange of memory contents.

TM Usage	never	occasionally	quite often	almost every job
(%)	57.2	16.8	8.1	17.9

Table 13: Use of Translation Memory Systems

Years of TM Use	0-2 years	3-5 years	6-10 years	more than 10 years
% (of TM-using respondents)	45.9	33.8	17.6	2.7

Table 14: Translation Memory Systems - Years of Use

Respondents were asked which TM systems they used, ranking from most commonly-used. Most used only one system.

TM System	Trados	SDLX	Wordfast	OmegaT	Deja Vu	TransAssist
System 1.	31	5	24	5	3	4
System 2.	5	2	1	1	2	-
System 3.	1	-	1	1	1	1

Table 15: Translation Memory Systems Used

Other TM systems mentioned included Metatexis, Wordfisher, Heartsome and StarTransit.

Reasons for TM Use	Required by Client	Expect productivity improvement	Experiment to see if it helps	Build memory for topic	Always use a TM
No. responding	38	45	20	35	21

Table 16: Reasons for Translation Memory System Use

Observed productivity improvement	less than 10%	10-25%	25-50%	over 50%
(%)	31.1	41.9	16.2	10.8

Table 17: Productivity Improvement with Translation Memory System Use

Client-supplied TM memory	no assistance	some assistance	moderate assistance	significant assistance
(%)	37.8	27.0	23.0	12.2

Table 18: Translation Memory Provision by Clients

TM Portability - aware of standards development	not at all	vaguely aware	recall the occasional article	following with considerable interest
(%)	73.4	11.6	9.8	5.2

Table 19: Translation Memory Portability – Awareness

TM Portability - will assist?	not at all	to a small extent	quite a bit	very significantly
(%)	65.9	13.9	13.3	6.9

Table 20: Translation Memory Portability - Assistance

Respondents were asked to list features they felt would significantly improve Translation Memory systems. Many listed basic aspects of TM packages, such as:

- lower prices and more transparent pricing structures;
- removing bugs;
- greater reliability;
- less complex user interface;
- Macintosh versions;
- better support for Microsoft's recent file types.

Desirable features mentioned included:

- easy generation of high-quality memories from existing translations;
- better glossary generation and term extraction;
- integration of online resources, including dictionaries and corpora;
- interchange of memories and glossaries;
- improvement of the "sentence" focus, to include working at both the phrase and paragraph level;
- improved collaborative working and server based memories.

8. Machine Translation Systems

The degree to which respondents were using MT systems, or working at editing machine translations was examined, along with perceptions of MT quality.

Use MT systems	Never	Occasionally	Regularly
(%)	91.7	7.1	1.2

Table 21: Use of Machine Translation Systems

Edit machine translations	Never	Occasionally	Regularly
(%)	83.4	15.4	1.2

Table 22: Edit Machine Translation System Output

Current MT quality	Don't know	Very poor	Marginal	quite good
(%)	29.0	58.0	13.0	0.0

Table 22: Current Machine Translation System Quality

Improved in recent years?	Don't know	No signs of improvement	Slight improvement	Getting significantly better
(%)	40.8	27.8	27.8	3.6

Table 23: Perception of Improved Machine Translation System Quality

Is MT a future threat to human translators?	never	in a few restricted areas	generally, but quality and rewriting still needed	we will all be displaced eventually
(%)	16.0	55.6	24.9	3.6

Table 24: Threat from Machine Translation Systems

Respondents were also asked to name up to three machine translation systems which worked into or out of Japanese. The names given were indicative of the issues of brand identity in commercial translation software. Babelfish was named 14 times, Alta Vista 4 times and Systran 4 times - in fact all of these use the same translation system (Systran), as did Google's translation function (6 mentions) until recently. The Fujitsu ATLAS system was mentioned 10 times, but no other commercial translation packages each received more than a single mention.

9. General Respondent Opinions and Expectations

Respondents were given the opportunity to express their opinions and expectations on various aspects of computers and translation. The categories in which view were sought were: technologies which have had the greatest impact on work as a translator; technologies which have not lived up to expectations; developments which could have a significant impact on translation work; and developments which could threaten work as a translator.

- a. technologies which have had the greatest impact on work as a translator.

A wide range of responses was received. Many reflected the particular circumstances of the respondent, with the more senior translators nominating "the computer", adding such things as "I started out using a typewriter". Clearly the Internet and related applications have had a huge impact on many aspects of translators' work.

The most frequently-mentioned technologies are summarized in the following table. (* Note that for many people the terms "WWW" and "Internet" are interchangeable.)

Technology/ facility	Internet(*)	WWW(*)	Email	WWW Search Engines	Online dictionaries & glossaries	Computers/ PC	TM systems
No. of responses	86	38	55	47	28	23	19
Technology/ facility	Word processors	Voice input	Online documents & databases	Wikipedia	OCR	GUI OS interfaces	FTP
No. of responses	13	10	8	4	4	4	4

Table 25: Technologies having an impact on translation

- b. technologies which have not lived up to expectations.

Not surprisingly, responses in this section largely focussed on machine/online translation (36 mentions) and translation memory/computer aided translation systems (29 and 7 respectively.) Views on translation memory systems are varied, with some respondents thinking they are improving, and others thinking they are not.

Of the remainder, voice input/recognition (8) and electronic dictionaries (7) were the highest, the latter receiving comments about lack of flexibility and integration. Others included collaborative translation management software (3) and Japanese OCR (3).

- c. developments which could have a significant impact on translation work.

A wide spread of responses was made in this section:

- i. translators expect to see faster and more ubiquitous communications facilities (14 mentions), and greater mobility in the working environment (9);
- ii. improvements in translation memory systems are expected and sought (25), with more use of artificial intelligence and statistical techniques being made;
- iii. machine translation systems are expected to improve (15), particularly in restricted areas such as patents. Several respondents expected to be working more in proof-reading/editing;
- iv. collaborative translation-support systems are expected to have an impact (13);
- v. speech input/voice recognition systems are expected to improve and be used more (9).

Other topics mentioned included: improved OCR, better online searching, more useful databases such as Eijiro and Wikipedia, more online texts, better online dictionaries, and greater English-language skill in client areas.

- d. developments which could threaten work as a translator.

Approximately half the respondents did not see any particular threat to their work. Of the listed threats, the most-mentioned was the rise of low-price competition, especially from India and China (32 mentions), although 3 respondents doubted its impact. Increased use of MT (18) and better TM systems (10) were the other commonly mentioned threats. Others mentioned included: global crisis/war (6), relative decline of Japan, a rise in English skills, more restrictive IP laws, and theft of intellectual property such as glossaries and memories.

10. Discussion

While the survey reported in this document cannot be regarded as wholly representative of the community of translators working with Japanese, as its distribution was primarily to those already working in online communities, it does provide valuable insights into the positions and opinions of a relatively large group of "computer savvy" translators.

In many ways the results confirm and reinforce the impressions formed by practitioners and observers of the translation profession that the impact of computer and communication technology has been dramatic over the last 15 years, with virtually every aspect of the industry seeing some change.

The author of the survey was mildly surprised at the relatively low level of translation memory usage reported, which when combined with the critical comments made by many respondents, indicates that for many translators either the benefits of translation memories are often overstated, or the hurdles to their use (price, training, complexity of interface, document interface problems, etc.) outweigh the advantages.

The comments made by the respondents throughout the survey are valuable, and indicate that the community surveyed is very alive to the issues associated with the business and technological environment of translation.

Acknowledgements

The author expresses his gratitude to the many translators who took the time to complete the questionnaire so frankly, and to Marc Adler, Ryan Ginstrom and James Sparks who reviewed the questionnaire prior to its release.