RESEARCH RESULTS DIGEST

June 2002—Number 267

Utility of Machine Translation Software

This digest summarizes the results of NCHRP Project 20-48 "Accessibility of Non-English Transportation Information." The digest is based on a report prepared by AppTek-L&H and Booz-Allen & Hamilton, Inc. The project objectives are to evaluate the usefulness of machine translation software in gaining at least a preliminary understanding of the contents of non-English language transportation documentation. The purpose of the study was not to evaluate or compare specific products, but to provide a general assessment of the current state of the art. Mohammad Shihadah served as the principal investigator.

SUMMARY

This digest provides an assessment of the current state of the art of software designed to translate foreign language documents into English. The study discusses the most appropriate use of machine translation and describes the level of intelligibility, accuracy, and style that can be expected from commercially available software. The information will be valuable for any transportation professionals who are looking for ways to access non-English information in their fields of expertise.

NCHRP Project 20-48 was initiated to investigate ways to improve access to non-English information for transportation professionals. One key task of the project was to benchmark machine translation (MT) software to determine its utility for translating non-English abstracts. The benchmarking activity focused on abstracts in the TRANSDOC database, as well as the other databases located at the University of California-Berkeley, Transportation Library. Four MT software packages were used to translate abstracts in French, German, and Spanish. The following conclusions were reached:

- There is a broad range of performance by commercially available MT software.
- MT software is suitable for "gisting," pretranslation, and, in many cases, aiding in human translation.
- The intelligibility, accuracy, and style of commercially available MT software fall substantially short of high-quality human translation. Translations were rated as vague or somewhat clear after one or two readings. Accuracy ratings reflected the fact that English translations typically had at least one important difference with respect to the non-English abstract. Style factors were often very different between the original and translated abstracts.
- A highly accurate lexicon of transportation terms can substantially improve the intelligibility and accuracy of translations so that, on average, they are intelligible after one reading and the meaning relationships between the original and source documents are fairly consistent.
- The state of the practice of commercially available MT software is constantly improving, and other available or recently released products already or will soon surpass the performance of the best MT software that was benchmarked.

Various lexicographical resources can be used to enhance machine-assisted translations. These include a Transportation Research Thesaurus recently developed for the Transportation Research Board and numerous transportation lexicons in various languages, including the World Road Congress (PIARC) multilingual lexicon, the European Community lexicon, and the European Conference of Ministers of Transport Lexicon.

BACKGROUND

Transportation practitioners and researchers often need to obtain transportation-related information from other countries. This information includes standards, specifications, guidelines, and manuals; details on best practices, research, and innovations; and publication lists. Although much

---

information is available in English, other important information is not. There is a pressing need for transportation professionals to be able to gain access to non-English materials.

Librarians and other information professionals work with a rich and complex transportation information infrastructure to respond to requests from practitioners and researchers regarding transportation-related information, including that not published in English. This information infrastructure includes online databases, the Library of Congress collections and catalogues, the World Catalogue (OCLC), the Internet, and listserves.

An important source of transportation-related information is the Transportation Research Information Service (TRIS), which contains more than 550,000 abstracts. TRIS is accessible through the Internet via the National Transportation Library, maintained by the Bureau of Transportation Statistics. The TRIS database contains only English language abstracts. Historically, the primary source of abstracts of non-English transportation-related documents has been the International Transport Research Documentation (ITRD) database, which contains abstracts in English, French, German, and Spanish. To date, there has been no simple, reasonably accurate, and cost-effective way to obtain translations of these abstracts and to archive and disseminate them. Similarly there has been no cost-effective way to translate, archive, and disseminate other more lengthy documents not in English.

Because of growing recognition of the need to improve access to non-English transportation-related information, NCHRP undertook a project to investigate improved ways to translate non-English transportation-related information and archive and disseminate translations.

As a result of advances in machine-assisted translation technology, a key objective of the project was to investigate the potential role of machine translation in providing the gist of documents not in English and to enhance full human translation.

BENCHMARKING STUDY

In this study, commercially available MT software was benchmarked to assess how well it could translate, into English, transportation abstracts in French, Spanish, and German. The benchmarking tests were conducted by Booz-Allen & Hamilton to provide an independent assessment of the MT software. Booz-Allen & Hamilton is not a developer or vendor of MT software and has no vested interest in a particular product. Software used to obtain each translation was not identified by vendor—only by an alphanumeric code.

EVALUATION OF THE MT SOFTWARE

Four different MT software packages were used to translate the French abstracts, four to translate the Spanish abstracts, and four to translate the German abstracts. Generally a range of commercially available software was selected and included software considered "best in class" or "state of the practice."

The results of the evaluation of each MT software package in translating abstracts are described in this section. Conclusions are drawn regarding the state of the practice of machine translation for each language and the applicability of the MT software for gisting and full translation of abstracts, as well as other potential uses.

Table 1 describes the functional, performance, and cost characteristics of the four French MT software packages. The software varies substantially in functionality and capabilities. The first three functional characteristics are rated on a scale of 1 (unacceptable) to 10 (excellent).

Ratings of the French MT Software

The rating methodology described above was used to assess the intelligibility, accuracy, and style of the translations of abstracts obtained using each French MT software. In all, 30 French abstracts (of generally no more than 250 words each) were translated. These abstracts comprised 158 sentences. Because MT3 was not able to translate 82 of these sentences, ratings were averaged only for the remaining 76 sentences.

The following issues arose in evaluating the machine translations:

- As mentioned previously, the French abstracts in the ITRD database did not include diacritics (i.e., accents) and these had to be inserted manually before applying the MT software.
- One of the software packages (French MT3) was unstable and caused the computer to crash and shut down. The software did not recover after repeated attempts to run it. Even with the aid of technical assistance from the vendor, the software would not work.
- Occasionally software would not translate an entire sentence. This was particularly true for French MT3.
- French MT4 produced mainly unintelligible translations because of very poor word substitutions.
- It became apparent, after completion of the evaluation of translations of the French abstracts, that the quality of the translation process might benefit from careful procedures to format the source material. This issue will be addressed when identifying appropriate technical procedures and business processes for machine translation in the second phase of the project.

TABLE 1 Functional, performance, or cost characteristics of French language MT packages

Functional, Performance, or Cost Characteristics	MT1	MT2	MT3	MT4
Robustness of software?	6	8	5	4
Ease of installation?	8	10	1	3
Ease of use?	10	8	5	5
Has general purpose dictionary?	Yes	Yes	Yes	Yes
Supports multiple specific domain dictionaries?	No	Yes	Yes	Yes
Able to update dictionary?	No	Yes	Yes	Yes
Allows user defined dictionary?	No	Yes	Yes	Yes
Can input glossary or terms?	No	Yes	Yes	Yes
Permits dictionary stacking?	No	Yes	Yes	No
Has translation memory?	No	No	No	No
Vendor support for upgrades?	NA	Yes	Yes	Yes
Vendor technical support?	No	Yes	Yes	Yes
Software cost?	$0*	$150	$135	$79

*Free 30-day examination copy over the Internet

Comparison of French MT Software

The average ratings for intelligibility, accuracy, and style, by all three raters for all sentences for each French MT software appear in Table 2 and in Figure 1.

Intelligibility

There is a broad range in intelligibility. The best software, MT2, achieved an average intelligibility rating of 1.50, signifying that the intelligibility of sentences translated from French to English falls between being vague and being somewhat clear after one or two readings. French sentences translated with MT1 produced an average intelligibility rating of about 0.89, indicating the average intelligibility was, at best, vague after one or two readings. Sentences translated with MT3 software were occasionally intelligible after one or two readings, but unintelligible most of the time. The worst software, MT4, produced mainly unintelligible translations.

Accuracy

Accuracy ratings of the translations of the French abstracts revealed a pattern similar to that revealed for intelligibility. MT2 software translations of sentences produced accuracy ratings of about 1 on average, the best score. MT2 sentence translations typically had at least one important difference between the original French and the resulting English, and often there was more than one important difference. Translations of sentences with MT1 software had accuracy ratings that averaged 0.62 as a result of being unintelligible much of the time. MT3 had an average accuracy rating of 0.4, and that of MT4 was nearly 0.

Style

The average ratings for the style of sentences translated from the French approximated the average ratings for intelligibility. MT2 had the highest style rating—about 1.35, indicating that the style factors were often very different

4

TABLE 2 Ratings for French MT software

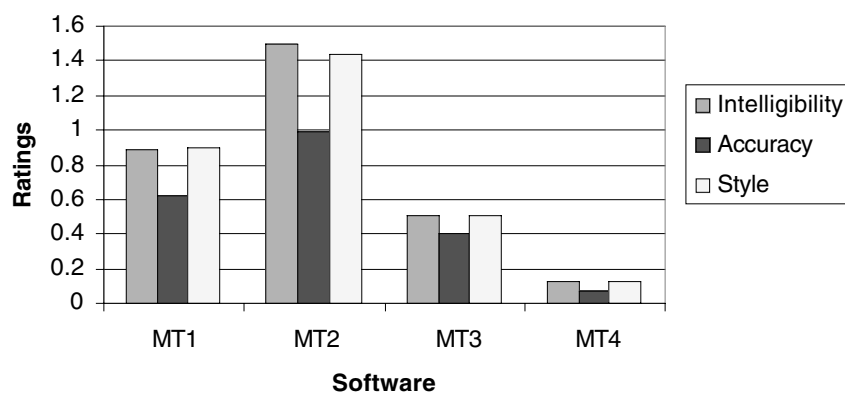|	MT1	MT2	MT3	MT4
Intelligibility	0.89	1.50	0.51	0.13
Accuracy	0.62	0.99	0.40	0.07
Style	0.90	1.44	0.51	0.13

Figure 1. Comparison of MT translations of French.

between the French original and the English translation, even though Spanish and English words were often cognates and the word order was often similar.

Ratings of the Spanish MT Software

The rating methodology described previously was used to assess the intelligibility, accuracy, and style of the translations of abstracts obtained using each Spanish MT software. In all, 29 Spanish abstracts (of generally no more than 300 words each) were translated. These abstracts comprised 200 sentences. Because MT1 was not able to translate 3 of these sentences, and MT3 was not able to translate 12 of these sentences, the average ratings for these software packages were calculated using the remaining 197 and 188 sentences, respectively.

The following issues arose in evaluating the machine translations:

- The Spanish abstracts in the ITRD database did not include diacritics (i.e., accent marks). The research team did not manually correct the text before applying the MT software in order to see the effect. In the best software, MT2, omission of diacritics seemed to have little effect, whereas in MT4, the effect on translation of specific words could be substantial.

- Occasionally, software would not translate an entire sentence. This was particularly true for Spanish MT3 and MT4.
- Spanish MT4 produced mainly unintelligible translations, probably because of very poor word substitutions.
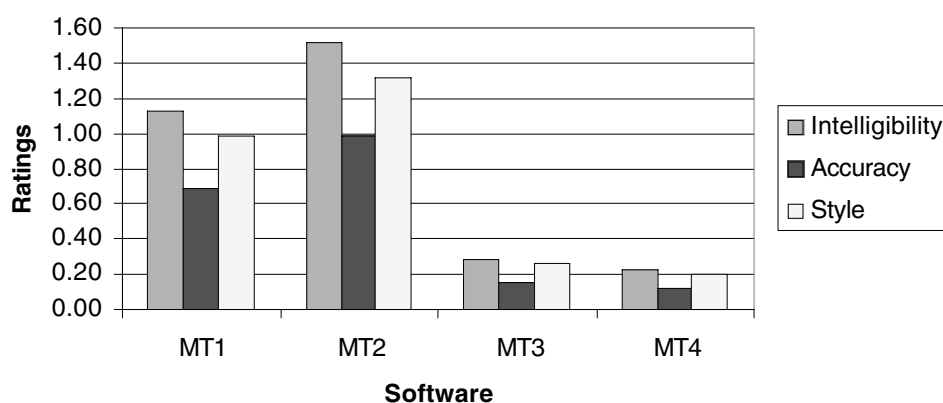
Comparison of Spanish MT Software

The average ratings for intelligibility, accuracy, and style, by all three raters for all sentences for each Spanish MT software appear in Table 3 and Figure 2.

Intelligibility

There is a broad range in intelligibility. The best software, MT2, achieved an average intelligibility rating of 1.52, signifying that the intelligibility of sentences translated from Spanish to English falls in between being vague or somewhat clear after one or two readings. Spanish sentences translated with MT1 produced an average intelligibility rating of about 1.13, indicating the average intelligibility per sentence was also vague or somewhat clear after one or two readings. Sentences translated with MT3 and MT4 software were unintelligible most of the time and seldom intelligible after one or two readings.

TABLE 3 Ratings for Spanish MT software

	MT1	MT2	MT3	MT4
Intelligibility	1.13	1.52	0.28	0.23
Accuracy	0.69	0.98	0.15	0.12
Style	0.98	1.32	0.26	0.20

Figure 2. Comparison of MT translations of Spanish.

Accuracy

Accuracy ratings of the translations of the Spanish revealed a pattern similar to that revealed for intelligibility. MT2 software translations of sentences produced accuracy ratings of about 1 on average, the best score. MT2 sentence translations typically had at least one important difference between the original Spanish and the resulting English, and often there was more than one important difference. Translations of sentences with MT1 software had accuracy ratings that averaged 0.69 as a result of being unintelligible much of the time. MT3 and MT4 had average accuracy ratings, indicating nearly total unreliability in the transfer of meaning from the source text to the target text.

Style

The average ratings for the style of sentences translated from the Spanish were slightly lower than the average ratings for intelligibility. MT2 had the highest style rating—about 1.32, indicating that the style factors were often very different between the original Spanish and the English translation, even though the Spanish and English words were often cognates and the word order was often similar.

Ratings of the German MT Software

The rating methodology described previously was used to assess the intelligibility, accuracy, and style of the translations of abstracts obtained using each German MT software. In all, 30 German abstracts (of generally no more than 300 words each) were translated. These abstracts comprised 236 sentences.

Two issues arose in evaluating the German machine translations:

- Occasionally, software would not translate an entire sentence. This was particularly true for German MT3.
- German MT4 produced unintelligible translations for the most part.

Comparison of German MT Software

The average ratings for intelligibility, accuracy, and style, by all three raters for all sentences for each German MT software appear in Table 4 and Figure 3.

Intelligibility

There is a broad range in intelligibility with the translations from German—this range is similar to those in the

TABLE 4 Ratings for German MT software

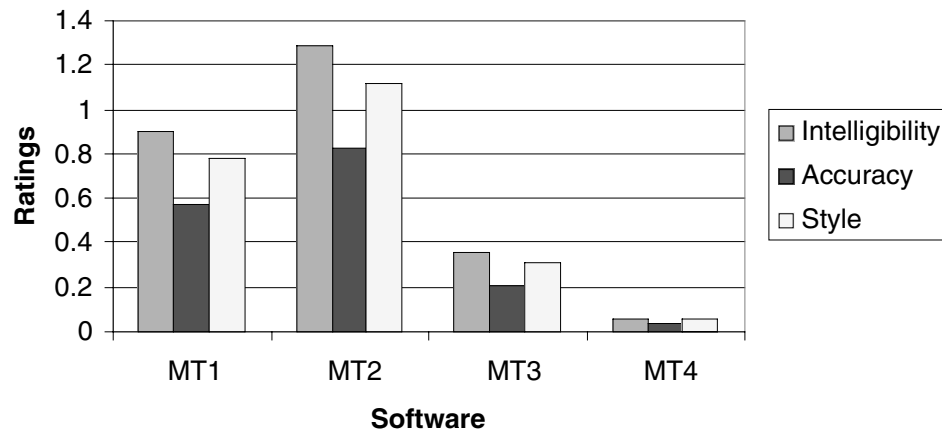|	MT1	MT2	MT3	MT4
Intelligibility	0.90	1.29	0.36	0.06
Accuracy	0.57	0.83	0.21	0.04
Style	0.78	1.12	0.31	0.06

Figure 3. Comparison of MT translations of German.

translations from Spanish and French. MT2 achieved the best average intelligibility rating of 1.29, signifying that the intelligibility of its sentences translated from German to English falls between being vague and being somewhat clear after one or two readings. Sentences translated with MT1 produced an average intelligibility rating of 0.90, indicating the average intelligibility per sentence was also vague or somewhat clear after one or two readings. Sentences translated with MT3 and MT4 software were seldom intelligible after one or two readings and were unintelligible most of the time.

Accuracy

Accuracy ratings of the translations of the German revealed a pattern similar to that revealed for intelligibility. MT2 software translations of sentences resulted in average accuracy ratings of 0.83 on a scale from 0 to 2. MT2 sentence translations typically had at least one important difference between the original German and the resulting English, and often there was more than one important difference. Even with the best software package, the meanings of the source and target texts were often completely dissimilar. Translations of sentences with MT1 software had accuracy ratings that averaged 0.57 as a result of being unintelligible much of the time.   MT3 and MT4 had average accuracy ratings, indicating nearly total unreliability in the transfer of meaning from the source text to the target text.

Style

The average ratings of each software translation from the German approximated the ratings for intelligibility. The ratings for style were the same or slightly lower than the ratings for intelligibility. MT2 had the highest style rating—1.12, indicating that the style factors frequently differ on at least one important level between the German original and the English translation, even though German and English words were often cognates and the word order was often similar.

FINDINGS ON COMPARISON OF MT SOFTWARE

These ratings suggest there is a broad range of performance by commercially available MT software. The test ratings indicate that the hierarchy for overall software quality, from best to worst, is MT2, MT1, MT3, and lastly MT4. None of the four software packages yield high-quality translations from French, Spanish, or German into English relative to human translations. However the best MT software is clearly suitable for initial accessibility and understanding of non-English transportation abstracts.

The benchmarking study suggests that MT2 represents the state of the practice of commercially available software. MT2 met the target ratings for gisting suitability. That is,

MT2 received intelligibility and style scores of at least 1, and the accuracy approximated 1. These scores would also be appropriate for a pre-translation exercise, where a user may be determining what types of additional dictionaries or lexicons might be needed in the next translation stage, be they automated or manual. MT2 is also likely to be useful in aiding in human translation, but a human translator cannot depend on the intelligibility, accuracy, or style of MT.

Although MT2 represents the state of the practice for commercially available MT software, other available or recently released products may surpass the performance of MT2.

EFFECTS OF ADDING TRANSPORTATION TERMS/LEXICONS

The purpose of the second part of the project was to assess the maximum intelligibility, accuracy, and style that could be achieved by incorporating a transportation lexicon into MT. Therefore, this performance assessment was carried out using MT2, which under the previous MT benchmarking test, was shown to provide the best translations.

Comparing the Results

For purposes of achieving the most accurate standard versus lexicon performance comparison possible, only the standard dictionary scores of the 10 abstracts in each language that were selected for translation by the "customized"

dictionary were tallied and averaged. The strategy was to average the translation ratings for each language lexicon added to MT2, and then compare these results with the computed averages for those offered using MT2 without the lexicon. Then an assessment was made to determine if there was a significant improvement in the average translation ratings for each of the three languages when a transportation lexicon was added to the standard dictionary in MT2.

German

The 10 German abstracts selected for this comparative analysis comprised 64 individual sentences/short paragraphs that each received the same type of ratings based on the three performance criteria used in the original MT analysis for the four base-product software packages. Table 5 lists the average translation ratings for the German MT2 standard dictionary and for the German transportation lexicon-supported dictionary.

According to these averaged translation ratings, the customized German transportation lexicon improved intelligibility by 37%, accuracy by 57%, and style by 25% for the same sentences and short paragraphs. Figure 4 further illustrates these improvements.

Spanish

The 10 Spanish abstracts selected for this comparative analysis comprised 50 individual sentences/short para-

TABLE 5 Ratings for MT2 German dictionary versus customized lexicon

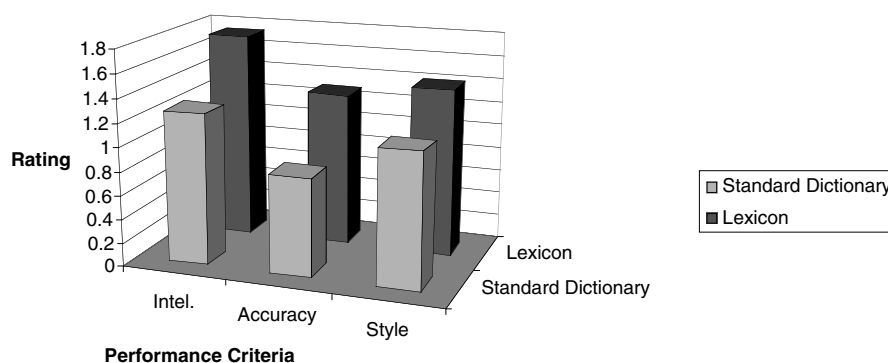|	MT2 Standard Dictionary	MT2 Standard Dictionary w/ Lexicon
Intelligibility	1.27	1.74
Accuracy	0.83	1.3
Style	1.13	1.41

Figure 4. German MT2 comparison of standard dictionary versus lexicon scores.

graphs. The average translation ratings for the Spanish MT2 standard dictionary and the Spanish transportation lexicon-supported dictionary are shown in Table 6.

The customized Spanish transportation lexicon improved intelligibility by 32%, accuracy by 54%, and style by 29% for the same sentences and short paragraphs. Figure 5 further illustrates these improvements.

French

The 10 French abstracts selected for this comparative analysis comprised 34 individual sentences/short paragraphs. Although this was the smallest sample set of 10 selected abstracts, the enhanced translation results conformed to those of German and Spanish. Table 7 lists the average translation ratings for the French MT2 standard dictionary and for the French transportation lexicon-supported dictionary.

These figures reflect that the customized French transportation lexicon improved intelligibility by 35%, accuracy by 33%, and style by 10% for the same sentences and short paragraphs. Figure 6 further illustrates these improvements.

FINDINGS REGARDING EFFECTS OF THE LEXICON

As a result of arming the standard MT2 dictionary for each of the three test languages with a customized transportation lexicon, the average increase in intelligibility was 35%. As expected for this exercise, scores for accuracy received the most dramatic improvement (48%). Style points increased by 21%—a significant increase for the most difficult and subjective of the three performance criteria. In general, the new lexicon-based translations read more smoothly and cleanly and, therefore, more easily than the translations produced by MT2 in the original analysis. Perhaps most importantly, for individual sentences and short paragraphs, the customized lexicons often provided the artificial intelligence necessary to make the difference between non-intelligibility and basic intelligibility in the translations to English.

It was evident through this process that, at least for MT2, adding a lexicon to a standard dictionary is a time-consuming process that requires expertise in the source language. It also became obvious that the quality of each cus-

TABLE 6 Ratings for MT2 Spanish dictionary versus customized lexicon

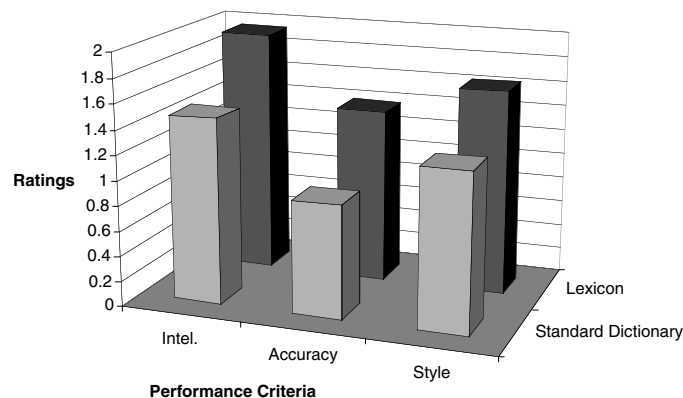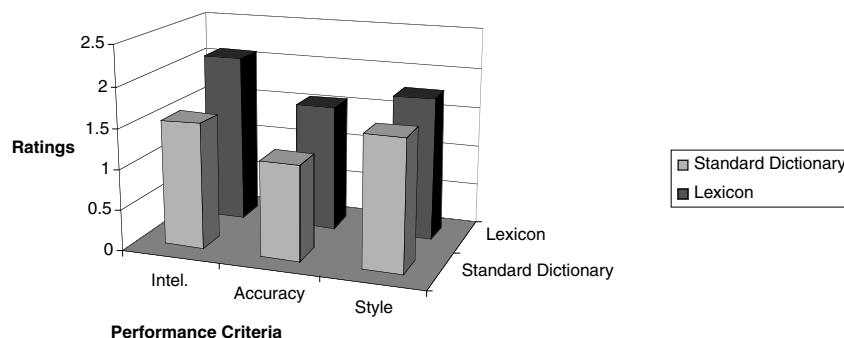|	MT2 Standard Dictionary	MT2 Standard Dictionary w/ Lexicon
Intelligibility	1.48	1.95
Accuracy	0.9	1.39
Style	1.26	1.63

Figure 5. Spanish MT2 comparison of standard dictionary versus lexicon scores.

TABLE 7 Ratings for MT2 French dictionary versus customized lexicon

Language	Abstract Cost-Low End	Abstract Cost-High End	Total Cost-Low End	Total Cost-High End
French	$30.20	$49.83	$431,467	$711,921
German	$32.20	$53.13	$1,748,820	2,884,533
Spanish	$40.00	$66.00	$457,200	$754,380
Total Cost			2,637,487	$ 4,350,834

Figure 6. French MT2 comparison of standard dictionary versus lexicon scores.

tomized lexicon was also influenced by the degree of expertise of the human interpretation of the source language. Standard and well-recognized transportation lexicons may or may not be useful depending on the source language and the transportation subtopic. These apparent realities of the user-friendliness of lexicons may be a great efficiency concern for English speakers who want to obtain large volumes of transportation-related information from non-English sources. However, other strategies to improve translation quality (e.g., incorporating "translation memory" of prior translated material) are likely to mitigate this problem and further enhance the quality of machine translation.

COSTS OF TRANSLATING ABSTRACTS IN THE ITRD DATABASE

The research team estimated the costs of translating the French, German, and Spanish abstracts in the ITRD database. Costs depend on such factors as the volume of material requiring translation, turnaround time, the need for technical domain expertise, and whether desktop publishing or special formatting is involved or not.

As of September 1999, there were 285,745 abstracts with approximately 1,000 being added monthly. The number and percentage breakdown by language appear in Table 8.

TABLE 8 ITRD abstracts and percentages

Language	Percentage of the ITRD	Number of Abstracts
English	72%	205,736
French	5%	14,287
German	19%	54,292
Spanish	4%	11,430

TABLE 9 Translation cost—performed by translation companies

Language	Abstract Cost-Low End	Abstract Cost-High End	Total Cost-Low End	Total Cost-High End
French	$30.20	$49.83	$431,467	$711,921
German	$32.20	$53.13	$1,748,820	2,884,533
Spanish	$40.00	$66.00	$457,200	$754,380
Total Cost			2,637,487	$ 4,350,834

TABLE 10 Translation cost—performed by freelance translators

Language	Abstract Cost-Low End	Abstract Cost-High End	Total Cost-Low End	Total Cost-High End
French	$15.10	$30.20	$215,733	$431,467
German	$16.10	$32.20	$874,410	$1,748,820
Spanish	$20.00	$40.00	$228,600	$457,200
Total Cost			$1,318,743	2,637,487

High- and low-cost estimates for translating these abstracts by translation companies and freelance translators appear in Tables 9 and 10.

The average number of words for abstracts was 151 words for French, 160 words for German, and 200 words for Spanish. Therefore, the cost varies per language.

ROLE OF TRANSPORTATION LEXICONS IN MACHINE TRANSLATION

The research team investigated the role that a bilingual or multilingual dictionary can play in improving machine translation. Experience in machine translation has shown that domain-specific dictionaries and dictionary stacking according to priority of context can substantially increase the quality of machine translation.

Survey respondents identified transportation lexicons in various languages, including the following:

- PIARC multilingual lexicon; three utilize the UIC (International Union of Railways) dictionary;
- The European Community Lexicon;
- The OECD (Organization for Economic Cooperation and Development) Lexicon;
- The ECMT (European Conference of Ministers of Transport) Lexicon;
- The TNC Lexicon (Swedish transportation lexicon); and
- The TransportordLista (Swedish transportation lexicon).

The Transportation Research Thesaurus prepared for TRB can also be used to develop a set of terms for which translations should be available in key languages.

An MT system is only as robust as its lexicon(s). Even if its morphological component, its syntactic parser, and its semantic analyzer are all based on the latest linguistic theories and represent state-of-the-art computational technology, all of this will be wasted if the system lacks a robust lexicon or set of lexicons.

Fortunately, if the substantive domain of the material to be machine-translated is limited, it is *not* necessary to construct a massive lexicon covering a myriad of subjects. One of the most successful, effective, longest-running MT systems in the world is the one used by the Canadian Weather Service to translate weather reports from English into French. Because weather reportage and prediction do not involve complex syntactic structures and because terminological domain is limited (e.g., "rain," "clouds," and "westerly winds,"), machine translation not only meets the challenge, but also has exceeded all expectations in producing high-quality output.

The situation with respect to the domain of transportation is somewhat analogous to that of weather reportage. Although a careful (human) reading of transportation-related abstracts reveals a high level of complexity and density in the prose, it also reveals the phenomenon of terminological consistency (i.e., the same terms cropping up time and again). If an MT system exists, which has proven its efficacy in other domains or in the general domain, it is safe to

conclude, therefore, that the same system could be brought to bear on the transportation domain, given the construction of a transportation-specific lexicon.

Furthermore, the use of electronic lexicons plays a significant role in the transportation community, aside from any consideration of machine translation. In the survey conducted for NCHRP Project 20-48, respondents reported that they either make active use of electronic lexicons or know of them (should the need arise to reference them). Three respondents reference the PIARC multilingual lexicon; three utilize the UIC dictionary. For each of the following terminology banks, one respondent reported utilization:

- The European Community Lexicon,
- The OECD Lexicon,
- The ECMT Lexicon,
- The TNC Lexicon, and
- The TransportordLista.

This indicates that the R&D investment in the construction of a transportation-specific multilingual dictionary could serve double-duty by (1) forming the lexical module for an MT system and (2) representing a stand-alone electronic lexicon for transportation professionals.

The MT products used in the evaluation process make use of the concept of "dictionary stacking": depending on the subject matter of the source document, the user decides which dictionaries should have reference-priority in the meaning selection process. For example, if the source document treats the topic of the air war over Kosovo, the user would "stack" the dictionaries in order of (1) military, (2) aviation, and (3) general purpose.

How and why does dictionary stacking improve MT output? Meaning selection differs in different domains. For example, the word "bug" may refer to an insect in the domain of entomology, while "bug" refers to a software flaw in the computer science domain. Dictionary stacking is not the only technique to aid in meaning selection. MT developers constantly face the problem of homograph disambiguation. For example, the word "turn" could be a verb meaning *to rotate* (etc.) or a noun meaning *a bend* (etc.). The syntactic parser should handle part of the job of disambiguation, determining the part of speech by contextual clues. Even after the part of speech is determined, however, the system may have to select among multiple senses.

Take the verb "break" for example:[1] *
(a) it could mean *announce*, as in "break the news";
(b) or it could mean *violate* as in "break the law";
(c) or *shatter* as in "break the glass";
(d) or *fail / breach* as in "break a promise";
(e) or *decipher* as in "break the code".

It is the job of the lexicographer to include *all* the meanings of each lexical item that he or she adds to the lexicon. A skilled lexicographer should be able to annotate each item for its meaning selection (semantics), its inflections (morphology), and how it fits with other words (syntax). Whenever possible, the lexicographer prefers to assign a single word meaning for each sense. At an average rate of between 2 and 3 minutes per lexical item, the lexicographer can add 200 words per 8-hour day. All of this assumes that the lexicographer already has in front of him/her the list of items (with their glosses) to be added to the lexicon.

A basis for developing a bilingual or multilingual transportation lexicon is the Thesaurus for Transportation Research (TRT). NCHRP Project 20-32, "Development of a Comprehensive Thesaurus for Transportation Research," and NCHRP Project 20-32(2), an extension of that project to extend the usefulness of the TRT, have culminated in the production of a machine-readable hierarchical thesaurus covering all modes of transportation and associated disciplines.

The TRT contains 11,627 postable terms, lead-in terms (UFs or "used-for" terms), scope notes, and container headings. Lead-in terms refer to items not considered to be U.S. standard terms, but which may be looked for by a user of the TRT (examples of such items would be British usage and spelling).

In extending the usefulness of the TRT, several micro thesauri have been developed, including those covering intelligent transportation systems, work zone safety, and traffic calming. A conversion dictionary was developed by the TRT team and has been used to convert terminology currently used in the TRIS database.

Project 20-48 is now completed. Copies of the contractor's final report, "Accessibility of Non-English Transportation Information," are available for a limited time upon request from the National Cooperative Highway Research Program, Transportation Research Board, 2101 Constitution Ave., N.W., Washington, D.C., 20418.

---

*This example is taken directly from AppTek-L&H's English-to-Arabic *TranSphere* MT product.