Translators at work with TRANSTYPE: Resource and Evaluation

Philippe Langlais, Marie Loranger, Guy Lapalme

RALI / DIRO
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada, H3C 3J7
http://www-rali.iro.umontreal.ca

Abstract

TRANSTYPE is an interactive machine translation prototype which has been designed at RALI with the hope that it could efficiently assist translators in their day work. During spring 2001, the latest version of our prototype was evaluated in an in situ setting. We first describe our evaluation protocol and then analyse the data we obtained. Second, we know that TRANSTYPE has been emulated and a few clones of it are already available. Therefore we have decided to open our log-files to the small but growing TRANSTYPE community. We believe such a database to be a helpful resource for both the researchers involved in the TRANSTYPE project, and also for all those interested in observing translators at work. Third, we briefly describe a free TRANSTYPE Player tool that we have developped; that is, a JAVA application which allows a given log-file to be played back.

1. Introduction

With few exceptions in very specific domains (*e.g.* (Chandioux, 1989)), and despite great improvements in machine translation (MT), it is a well known fact that translators feel reluctant to post-edit the output of a machine translator. TRANSTYPE originated in the mid-nineties (Foster et al., 1997) with the idea that automatic translation technology could nevertheless be of use if conceived in an interactive setting.

Currently, TRANSTYPE takes the form of a text-editor which embeds a probabilistic engine whose role is to provide at any time a pop up menu containing a ranked list of completions that may extend what the translator is typing. Thus, a translation emerges from alternating contributions by human and machine, with the translator's inputs serving as progressively informative constraints for the MT component (see Figure 1).

It has been shown that under a word-completion scenario, a user who carefully looks at the proposals made by TRANSTYPE and accepts them on purpose may save around two thirds of the keystrokes needed to type the full text (Foster et al., 1997). During spring 2000, a mature prototype was evaluated in an *in situ* setting. Ten translators came to our laboratory and used TRANSTYPE for more than one hour each. The results of this study have been fully described in (Langlais et al., 2000).

Several small mistakes were made during this evaluation that slightly polluted the interpretation of the figures we observed. For example, we made a few bad ergonomic choices that rendered the use of TRANSTYPE less obvious than we thought. We also did not realize that timing each interaction to the millisecond was necessary to make clear measurements of some tricky observations.

Among the qualitative feedback we gained from this evaluation, translators mentioned how important lexicons are in the translation process. Actually, lexicons are probably the only means a user has to tune a translation engine to his/her special needs. Therefore, we decided to extend TRANSTYPE so that it now allows for the on-line integration of user lexicons. The technical details behind this new implementation are fully described in (Langlais et al., 2001).

For all these reasons, we decided to carry out a new *in situ* evaluation of TRANSTYPE during summer 2001. Section 2. provides a description of the evaluation protocol we applied. In section 3., we analyse the log-files we collected. In section 4. we describe a resource freely available on the Internet which offers in a ready-to-use format all the interactions between TRANSTYPE and its users. This resource comes with a graphical interface (a java applet) which allows a given session to be played back.

2. The summer 2001 evaluation protocol

An *in situ* evaluation of TRANSTYPE was carried out during spring 2000; and the results of this study have been described in (Langlais et al., 2000). Since then, the prototype has been improved and it is now able to suggest completions that go beyond the word level (see Figure 1 for such a completion). This is done by allowing the user to incorporate at any time his or her lexicon that may contain units associations (*e.g. shipbuilding*/construction navale).

At the same time, we corrected a few ergonomic choices that were made in the previous version of the prototype and which were responsible of many misuses of the prototype. Among these, in the previous version of TRANSTYPE, the ENTER key was used to move to the next source sentence to be translated (this choice was made at a time when TRANSTYPE did not have a graphical interface). But we observed that the inclination of many users was to use this key to accept the top menu completion. Therefore, it often happened that a user unexpectedly changed the sentence being translated, thus losing some time to come back to the previous sentence to pursue its translation.

With the goal of evaluating TRANSTYPE in its new setting, we designed a new evaluation protocol that required about one hour of a user's time.

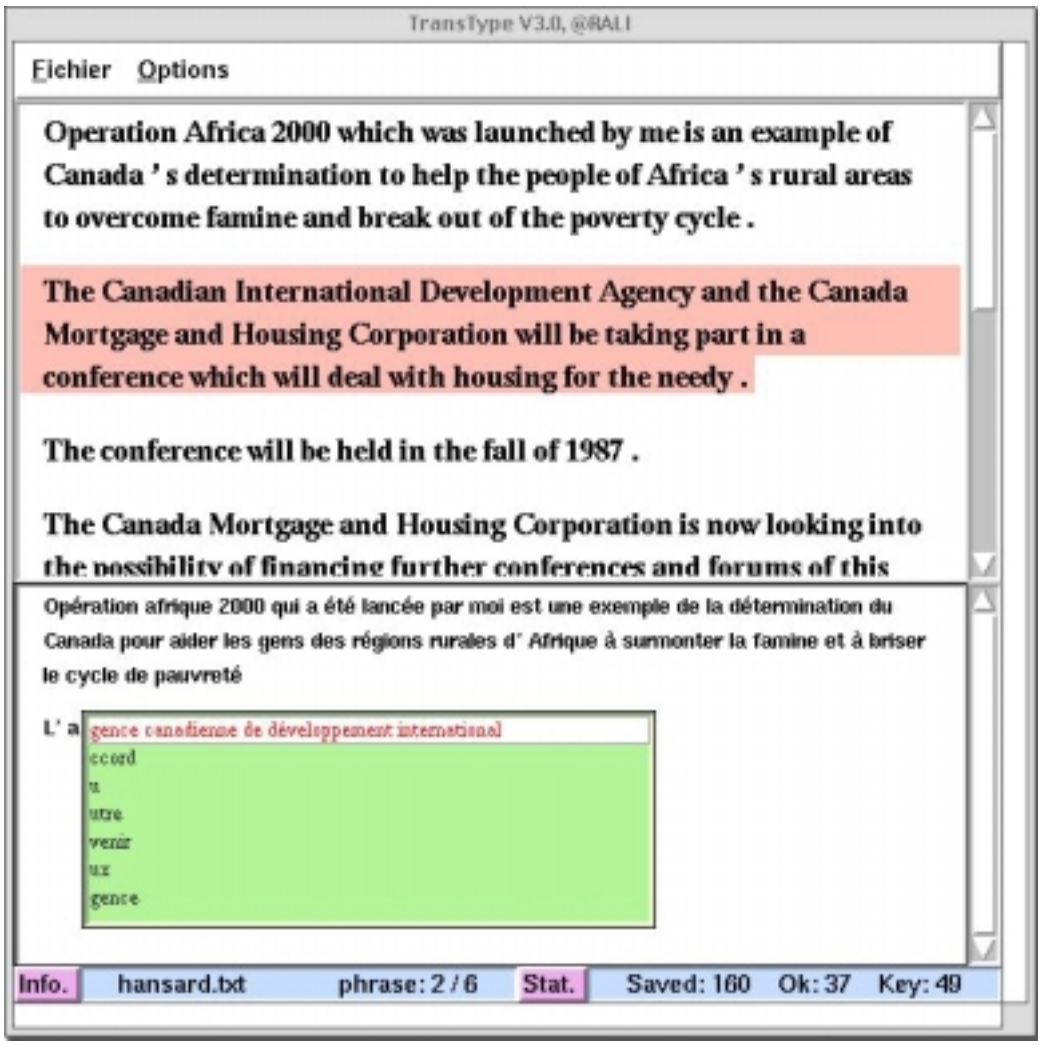

Figure 1: An example of interaction in TRANSTYPE with the source text in the top half of the screen. The target text is typed in the bottom half with suggestions given by the menu at the insertion point.

2.1. Road map of the evaluation protocol

Introduction The user is first introduced to TRANSTYPE and to the goal of the evaluation. General instructions are given during this introduction. In particular, we emphasize that the translator should not worry about formatting matters, but instead should focus on producing a version whose content would require a normal review. We must stress here, that it is not our goal to grade the quality of the translation produced, although we think it would be very interesting to know whether TRANSTYPE has any impact (positive or not) on translation quality (see the qualitative survey for more on that question).

Step 1 (5-8 minutes) The first step of the protocol puts the subject directly in contact with the text-editor implemented in TRANSTYPE. This editor offers all the standard operations (cut & paste, delete, etc.) that a person familiar with computers may expect. During this period, TRANSTYPE works in a silent mode (i.e. it does not propose anything) and the user only uses the editing functionalities of the prototype. We expect that this 5-8 period stage will make the user familiar with the few specific commands TRANSTYPE requires (e.g. selecting a new source sentence to translate). We also used this stage to calibrate the user's typing and/or translating speed.

Step 2 (15-20 minutes) In the second step, TRANSTYPE is switched to its normal mode, that is, proposing after each keystroke the completion of the current word, or the current sequence of words whenever it belongs to a general purpose lexicon automatically integrated into TRANSTYPE.

Step 3 (5-8 minutes) The purpose of this phase was to gauge the usefulness of the new functionality we added to TRANSTYPE; that is, integrating a user lexicon; which to begin with was manually designed for the special text we asked the users to translate here.

Survey The evaluation protocol ends up with a feed-back survey to collect the subject's feelings and suggestions.

Before steps 1 and 2, some time was alloted to the translator to become familiar with the assignment. Over a period of at most 5 minutes, the user was allowed to try, without being logged, to become familiar with the assignment. This was done in the hope of reducing the slow down we usually observed at the very beginning of each step, due to hesitations and ergonomic adaptations.

No other resource than TRANSTYPE was available to the user at the time of translating. In particular, and despite the fact that some users were disturbed by this, we did not provide them with a dictionary. Consultation of a dictionary would have made the timing of our experiments difficult to automatise.

2.2. Material used during this evaluation

The material used in the two first steps consisted of a corpus of about one hundred isolated sentences chosen from the Hansard corpus. We excluded the sentences that had been used during the training of the language and the translation models and also removed sentences that were too long, contained too many complicated proper names or numbers, etc. Finally, we inspected the selected sentences in order to remove those that we found to be ambiguous or difficult to translate without larger context (e.g. sentences with ellipses, etc).

During step 2, it was sometime the case that sequences of words were proposed instead of just a single word. This was due to the fact that TRANSTYPE automatically integrates a general purpose lexicon. Actually this lexicon was automatically acquired from an excerpt of the Hansard corpus. Table 1 gives some of the entries belonging to this lexicon.

source entry	target entry
• family allowance cheque	• chèque d'allocations familiales
• export enhancement program	• programme de stimulation des exportations • programme d'améliorations des exportations
• federal cuts	• compressions budgétaires fédérales
• senior officials	• hauts fonctionnaires

Table 1: A few entries from the general purpose lexicon automatically integrated in TRANSTYPE.

For step 3 (the one with the user lexicon), we took an excerpt of a text from the Health Canada Internet site: "Nutrition for Healthy Term Infants". This text is by nature fairly different from the ones we trained our translation engine on. We think it would reflect more truly the environment in which TRANSTYPE may be used. We also think that in such cases, the need for specialized lexicons would also be stronger. The lexicon we integrated in TRANSTYPE during step 3 is composed of 14 entries that are given in table 2.

3. Analysis

The goal of this evaluation was two-fold: to measure if TRANSTYPE saves time for its users and to gauge if it helps in other ways, such as giving ideas for the translation of terms for which there is some hesitation. As the suggestions of TRANSTYPE are correctly spelled, their selection reduces the number of misspellings in the target text. This is particularly useful for completed proper nouns or numbers, which must always be carefully transcribed and are often prone to error.

3.1. The qualitative survey

In our two sets of experiments, all our subjects (except one) were enthusiastic about this concept of a translation typing tool even though our prototype was far from being perfect. At the end of the session, the users were invited to answer a few questions and to comment on the prototype they used. Here are the answers to some questions we asked:

- *Do you feel* TRANSTYPE *makes you faster ?*

  4 answered "no", 1 said "not at the beginning", 1 said "maybe", two said "yes".

  It is worth mentioning that none of these users actually managed to go faster using the completions proposed by TRANSTYPE.

- *What do you think of the possibility of adding your own lexicon within* TRANSTYPE*?*

  All the users said it is really a must. For four of them, it would even justify the use of TRANSTYPE in their day work.

- *Do you think* TRANSTYPE *had an impact on the quality of your translation ?*

  Most of the users had the feeling that TRANSTYPE had a negative impact on their translation. Interruption of the mental process and the tendency to translate literally are some of the problems mentioned. However, users appreciated however the completion of difficult entities, such as numbers and proper names.

- *Do you have any suggestions to improve* TRANSTYPE*?*

  The main ones are: a) remove short completions from the pop-up menu; b) integrate TRANSTYPE within a real text editor; c) propose only the baseform of the words.

Most of the users were confident that with time they would become more proficient at making better use of TRANSTYPE. One translator pointed out that once he had an idea, he went along without looking at the suggestions; and in fact, he almost never used the tool. We were glad to observe that the users really appreciated the possibility of adding their own lexicon. We know, however, that this kind customization is always appreciated because users like to make their tool their own, even though in practice very few take the time to really adapt their existing tools, such as their text editor.

source entry	target entry
healthy term infants	nourrissons nés à terme et en santé
dietitians of canada	les diététistes du canada
health canada	santé canada
public health	santé publique
partly skimmed milk	lait partiellement écrémé
skim milk	lait écrémé
breastfed	nourri au sein
breast milk	lait maternel
canadian paediatric society	société canadienne de pédiatrie
commercial iron-fortified formulas	les préparations lactées commerciale enrichies de fer
infant formulas	préparations pour nourrissons
modes of feeding	modes d'alimentation
canada food guide to healthy eating	guide alimentaire canadien
iron deficiency	carence en fer

Table 2: The specialized lexicon manually designed for step 3 of our protocol.

3.2. Quantitative analysis

A theoretical evaluation of the translation engine (Foster et al., 1997) has shown that TRANSTYPE can save about two thirds of the keystrokes needed to type a given translation (at least in situations where the text to be translated is close enough to those used at training time). Unfortunately, the results given in table 3 are quite different. The users saved on average only 31% because they did not use all the completions that were available to them. From the study of the logs keeping track of the users' actions, we infered that users could have saved up to 68% of the keystrokes if they had always chosen the best completions that were available to them.

subject	typed	accept	erased	final	% typed
1	1528	1001	181	2348	60%
2	791	411	181	972	66%
3	1234	582	149	1667	68%
4	1255	164	60	1360	88%
5	554	311	77	789	64%
6	1220	757	191	1786	62%
7	374	260	94	537	59%
8	634	352	178	809	64%
9	2198	332	407	2123	87%
2001	1088	463	169	1377	69%
2000	486	972	111	1347	55%

Table 3: Number of characters typed, accepted by validating the suggestions of TRANSTYPE, and erased. The fifth column reports the number of characters present in the text produced at Step 2 of our protocol. The last column shows the proportion of characters manually typed over the number of characters in the final text. The last two lines indicate the mean for this year's (2001) evaluation and for last year's (2000).

These results are noticeably different from those of last year, where users had saved 45% of the keystrokes (see the last line of table 3).

When a user accepted a completion, it was 23% of the time with the mouse, 38% with the return key after cycling through the list an average of 3.1 times, and 39% using only the return key because the completion was the first choice.

3.2.1. Productivity

We define productivity as the ratio of the number of characters in the final text over the time it took to produce the text. In the following, we express this quantity in terms of the number of characters produced in a minute. In interviews, some of the translators revealed that they thought that TRANSTYPE had improved their productivity. Unfortunately, Table 4 does not corroborate this favorable impression, because on the average, raw productivity actually went down by -17%, an improvement over last year's -35%!

Subject	Step 1	Step 2	Gain	Step 3	Gain
1	153	112	-27 %	89	-41 %
2	28	53	88 %	43	53 %
3	114	94	-17 %	63	-45 %
4	86	84	-3 %	79	-9 %
5	79	48	-39 %	57	-32 %
6	113	104	-7 %	58	-45 %
7	53	37	-30 %	61	-7 %
8	64	40	-38 %	43	-32 %
9	145	119	-18 %	137	-17 %
2001	93	77	-17%	70	-25%
2000	102	65	-35 %		

Table 4: The productivity of the translators at each step of the protocol. Gain figures are computed with the productivity measured in step 1. The last line indicates the mean for all translators. The last two columns give the productivity and gain for step 2 but not counting the first 10 minutes in order to take into account the learning process of the interaction with TRANSTYPE.

One subject (no 2) who was especially slow in the first step made a very productive use of completions. With the lexicon, users still lost more productivity; we will come back to this point later.

Questions can be raised about the learning curve for such an "unusual" tool as TRANSTYPE. We did some statistics by ignoring what had been done in the first ten

minutes of Step 2. Of course, statistics are then computed over a much shorter time (5 to 8 minutes), but we can still observe an important difference (-10% lost instead of -17%). This saving does not seem to depend on the speed of the user. The proportion of manually entered characters is still the same at about 70%.

These results are only preliminary and, as confirmed verbally by our subjects, the use of TRANSTYPE over a longer period would give a much better picture of its possibilities. We are inclined to think that after ten minutes, users were more concentrated on their text and less distracted by the context of the experiment and the novelty of the tool.

3.2.2. Behavior with respect to suggestions

The analysis of suggestions proposed by TRANSTYPE incurs a certain "cognitive load" on the user because of the interruption in the translation process. On average, there is 0.75s time-lag between the time a suggestion is proposed by TRANSTYPE and the next action from the user. But when the user accepts a suggestion, the average goes up to 1.47s, almost doubling the whole average. When a suggestion is not accepted, a user reacts in less than 0.3s half of the time, which is quite short compared to a 0.65s average between the keys of two successive characters measured in the first step. One can thus assume that users did not stop to look at the propositions or did not see them. To understand when TRANSTYPE should display its suggestion, it is interesting to measure after how many characters on average a user accepts a suggestion. In step 2, 1535 units (71%) were typed by the users, 311 (14%) were entered by a completion spanning the whole word, and 100 (5%) were completed after typing only one letter. Less than 10% of the words were completed after two or more characters had been entered.

Looking at the reaction time and the length of the completions, we can conclude that when a user accepts a completion, it is at the beginning most of the time. Thus the best way to improve the use of the suggestions would be to convince translators that they should look at the suggestions very soon in their typing process. As looking at these suggestions and deciding if they are worthy takes time and can in a way distract from the thinking process of the user, suggestions must be valuable i.e. long enough.

3.2.3. Length of useful suggestions

As we have been told by translators, it seems that suggestions of 3 letters or less are not very useful. We had essentially similar comments in our previous evaluation but this year we can corroborate this impression with "hard" figures by looking at the average time between two keystrokes (0.65s) and between the display of a suggestion and its approval (1.4s). This means that a user can type more than two characters in the time it takes to read and accept a suggestion. So with a suggestion of less than three characters, the user is bound to loose time. Hence, one lesson is that suggestions of less than three (and probably less than four) characters should not even be displayed.

In our experiment, the average length of accepted completions was 5.5 characters, but 42% of these were less than four characters (51% of the displayed suggestions had less than four characters). Given the fact that these short completions are likely to incur some loss in productivity, we can in part explain why translators have been slower with TRANSTYPE (Step 2) than without (Step 1).

3.2.4. Use of the lexicon

In the third step of our experiment, we decided to work on a text for which we built an appropriate lexicon by hand (see Figure 2).

In this step, accepted completions were longer (7.5 characters instead of 5.5) but the loss in productivity was larger than for Step 2 (see table 4). The different domain of discourse surely had an important effect on this result. We also observed that translators often started the translation of a unit differently from the terms we had inserted in the lexicon and thus TRANSTYPE could not propose an appropriate completion from its lexicon. We suspect this would not have been observed if the user himself had compiled the lexicon.

Although the numerical results do not show it clearly, interviews with the users revealed that customization with a lexicon is very appreciated, as we have discussed in thea subsection 3.1.

3.2.5. Performance of the translation engine

A great deal of effort in our project has been devoted to the development of the statistical translation engine, so it is interesting to examine its performance in a real setting. In step 2, TRANSTYPE suggested an appropriate translation for 899 words (42%) that were accepted by the translator, and for 747 of them (35%) the first suggestion was the right one. This means that TRANSTYPE's suggestions were "optimal" 77% of the time. Only 376 words (17%) did not get any completions and about half of these were words with spelling errors or apostrophes. Due to a bug in our prototype, TRANSTYPE could not deal appropriately in this experiment with words containing apostrophes.

4. A free database of translators at work

We presented the first *in situ* evaluation of TRANSTYPE at the LREC'2000 conference and there it was suggested that we make the log file of the translation sessions available. Although we constantly postponed the work of transforming all our log files into a friendly enough format, this suggestion slowly matured. When we decided to carry out the second evaluation of TRANSTYPE, we designed a new format to trace all interactions between the user and the computer. This format is now clearly documented and allows us to open our log-files to the community.

We see several reasons why this resource may interest others. First of all, the first stage of the protocol (the one in which TRANSTYPE is silent) provides materials that should reflect how translators work. Considering that all interactions were timed, this allows for some interesting studies related to the translation process.

Second, on March 1st of this year a new project called TT2 (for TRANSTYPE2) was officially launched, founded by the European Council of Research. This means that clones of the current TRANSTYPE prototype will soon appear within the community. In fact, at RALI, we already have several versions of TRANSTYPE. Providing a

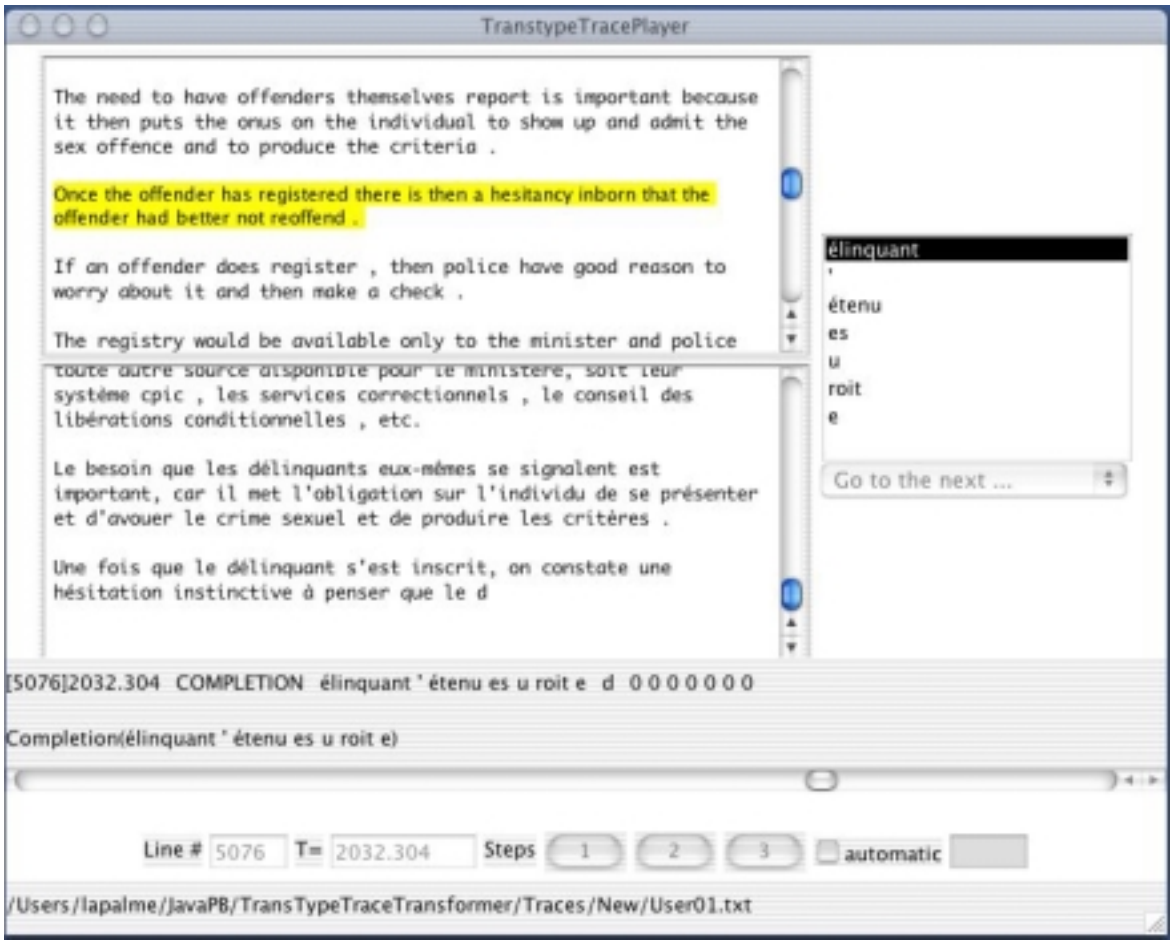

Figure 2: Screen dump of the log-file player

TRANSTYPE database of log files in a well described format would therefore be of help.

4.1. The TRANSTYPE database: TTBase

Nine sessions are currently available on the TRANSTYPE's web page[1]. This represents a total of 243 sentences produced during steps 2 and 3. This seems few compared to the megabytes of text current tools nowadays process, but we do feel, considering its very specific nature, that this represents a valuable resource that we encourage the community to take advantage of.

Because of the simplicity of the information available in our log-files, we decided to distribute them almost verbatim (only the identity of the translators has been removed from the original log-files). Basically, any action that occurred during a session is logged with its precise time.

There are two kinds of actions logged in this database: *environmental* ones, and *editing* ones. The former are actions that are more concerned with the protocol we described (starting a given evaluation session, a given step of the protocol, inserting a user lexicon into the engine, etc). The latter are directly logging the interactions between the user and the engine to produce the translation (inserting/removing a character, accepting a suggestion, cycling through the pop-up menu, etc.).

The full description of the format is provided in a document available on the TRANSTYPE's project web page. Figure 3 provides a small excerpt of such a log file. As can be seen, our format has the advantage of being ready to use. Actually, data reduction may be carried out very easily by script-like commands, or by direct manipulation within a spreadsheet such as Excel.

Last but not least, we also provide a TRANSTYPE player, that is a JAVA applet which allows a given session to be played back from a log-file. This may be helpful for instance to understand why a good statistical engine is not necessarily the one that makes the user faster. See Figure 3.2.4. for a screen dump of this tool.

5. Discussion

The contribution of this work is twofold: first, we have described the latest evaluation we carried out of our current prototype. This allowed us to measure precisely that of the some ergonomic choices implemented in our prototype are not well motivated. Among these, we concluded that proposing very short words (currently, at least 42% of the suggestions made by TRANSTYPE in this experiment) is a counter-productive way of interacting with the user.

Second, we invite the community to take advantage of a new kind of database which is available at the address `www.iro.umontreal.ca/TransTypeProject-`

[1] http://www-rali.iro.umontreal.ca/ProjetTransType.en.html

en.html. This should be useful to anybody interested in TRANSTYPE, or more generally in any target text mediated interactive machine translation system. It should also be of benefit to anyone interested in the study of the translation process itself. The data reduction we described in section 3. is one example of how this database can be used.

There are many ways to improve our current prototype. Better models would of course be a plus. The current prototype makes use of a linear combination of an interpolated trigram target language model (Jelinek, 1990) and an IBM2-like source-to-target translation model (Brown et al., 1993)[2]. The embedded decoder is the most simple one imaginable: pick the 7 best words according to this combination. Foster (2000) has investigated a maximum entropy model where both a trigram and an IBM2-like models are combined in a more principled way. This model is much more compact than the one currently used, and it has been shown that it can significantly improve the performance of TRANSTYPE.

Very recently, we investigated the possibility of embedding a viterbi-like decoder within TRANSTYPE. The problem is not as simple as one may think. First, stringent time constraints inherent to the approach may give classical beam searches a hard time. Second, we cannot realistically overwhelm the user with too many long suggestions, as it takes time to process them. The preliminary experiments we have conducted indicate that controlled completions of up to 5 words may be fast and accurate enough to improve the overall performance of TRANSTYPE. The results od this study will be described elsewhere.

Acknowledgments

TRANSTYPE is a project which has been funded by the Natural Sciences and Engineering Research Council of Canada. We are indebted to George Foster and Elliott Macklovitch for the fruitful orientation they have given to this work. We also want to warmly thank the subjects who participated in this study and from whom we have gained much insight.

6. References

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. 1993. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19(2):263–311.

J. Chandioux. 1989. Météo: 100 million words later. In D.L. Hammond (ed.), editor, *American Translators Association Conference: Coming of Age, Learned Information*, pages 449–453, Medford NJ.

George Foster, Pierre Isabelle, and Pierre Plamondon. 1997. Target-text Mediated Interactive Machine Translation. *Machine Translation*, 12:175–194.

George Foster. 2000. A Maximum Entropy / Minimum Divergence translation model. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 37–44, Hong Kong, October.

[2]For an intensive comparison of several translation models, see (Och and Ney, 2000); for an alternative translation model see for instance (Yamada and Knight, 2001).

F. Jelinek. 1990. Self-organized language modeling for speech recognition. In A. Waibel and K. Lee, editors, *Readings in Speech Recognition*, pages 450–506. Morgan Kaufmann, San Mateo, California.

Ph. Langlais, S. Sauvé, G. Foster, E. Macklovith, and G. Lapalme. 2000. Evaluation of transtype, a computer-aided translation typing system: A comparison of a theoretical- and a user- oriented evaluation procedures. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC)*, pages 641–648, Athens, Greece, June.

Philippe Langlais, George Foster, and Guy Lapalme. 2001. Integrating bilingual lexicons in a probabilistic translation assistant. In *Proceedings of the 8th Machine Translation Summit*, pages 197–202, Santiago de Compostela, Galicia, Spain, September.

F.J. Och and H. Ney. 2000. A comparison of alignement models for statistical machine translation. In *Proceedings of the International Conference on Computational Linguistics (COLING) 2000*, pages 1086–1090, Saarbrucken, Luxembourg, Nancy, August.

Kenji Yamada and Kevin Knight. 2001. A syntax-based statistical translation model. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 523–530, Toulouse, France, July.

```
2491.612   COMPLETION   le de et les qui {les diététistes du canada} la
2493.593   NEXT_COMPLETION de    0  0.0258675414645
2493.595   COMPLETION   de et les qui {les diététistes du canada} la le
2493.766   NEXT_COMPLETION et    0  0.0199645940585
2493.769   COMPLETION   et les qui {les diététistes du canada} la le de
2493.912   NEXT_COMPLETION les     0  0.0164671885459
2493.915   COMPLETION   les qui {les diététistes du canada} la le de et
2494.334   NEXT_COMPLETION qui     0  0.0161261647843
2494.337   COMPLETION   qui {les diététistes du canada} la le de et les
2494.500   NEXT_COMPLETION les diététistes du canada     1  0.0158702722952
2494.502   COMPLETION   {les diététistes du canada} la le de et les qui
2494.928   KEY_ACCEPTED_COMPLETION les diététistes du canada     1 0.158702722952
2494.930   PASSIVE_INSERTION
2494.954   COMPLETION   et a est à de en qui
2497.435   KEY_ACCEPTED_COMPLETION et    0  0.0449701775849
2497.438   PASSIVE_INSERTION
2497.463   COMPLETION   les {les diététistes du canada} de le la {la santé} à
2501.077   ADD_CHAR   S
2501.084   COMPLETION   es on ur {ociété canadienne de pédiatrie} ociété {anté canada} anté   S
2501.208   ADD_CHAR   a
2501.211   COMPLETION   {nté canada} nté ns voir nitaires nitaire  Sa
2501.393   ADD_CHAR   n
2501.396   COMPLETION   {té canada} té s itaires itaire San
2501.570   ADD_CHAR   t
2501.573   COMPLETION   {é canada} é  Sant
2501.734   ADD_CHAR   é
2501.736   COMPLETION   { canada}  Santé   1
2501.873   ADD_CHAR
2501.898   COMPLETION   et nationale ont des de au du
2503.008   ADD_CHAR   c
2503.012   COMPLETION   anadien anadiens e anadienne omme ette ’   c
2503.178   ADD_CHAR   a
2503.181   COMPLETION   nadien nadiens nadienne nadiennes r s use  ca
2503.254   ADD_CHAR   n
2503.257   COMPLETION   adien adiens adienne adiennes adian   can
2503.384   ADD_CHAR   a
2503.387   COMPLETION   dien diens dienne diennes dian  cana
2503.559   ADD_CHAR   d
2503.562   COMPLETION   ien iens ienne iennes ian  canad
2503.704   ADD_CHAR   a
```

Figure 3: Excerpt of a log-file from a user translating the sentence: *Prepared by the Canadian Paediatric Society , Dietitians of Canada and Health Canada.* The user has already typed the text (not shown here) Société canadienne de pédiatrie, (by accepting a unit completion). The first column reports the time of action (expressed in milliseconds spent since the beginning of the session). Here is the sketch of how the translation went: a) the user cycled through the pop-up menu to get the suggestion les diététistes du canada (time *2491.612* to *2494.930*); b) s/he accepted the completion et (time *2494.954*) and c), s/he typed the string Santé Canada without considering the suggestions made by TRANSTYPE. This last string was proposed at time *2501.084*.