TransCheck –
or the Automatic Validation of Human Translations

Elliott Macklovitch[1]
Centre for Information Technology Innovation
Montreal, Canada

1. Introduction

Translators, like all people who draft a lot of texts, will commonly run a spelling or style checker on their translations before delivering them, in the hope of detecting simple kinds of errors they may have overlooked. Because these writing aids are monolingual, however, they are entirely useless in detecting even the most flagrant kinds of *translation* errors, since these are *bilingual* in nature and depend on relations between two texts in different languages. This paper will describe a novel kind of writing aid called TransCheck that is specifically designed to detect such errors.

TransCheck is one of a family of new translator support tools being developed at the CITI (see reference [5]), all of which are based on alignment algorithms like those originally proposed by Gale and Church [2] and Brown et al. [3] for parallel texts. What these alignment algorithms do, in a word, is to calculate the links between corresponding segments in a source text and its translation, with a high degree of accuracy.[2] Following Harris [4], a corpus of translated texts aligned in this way may be called a *bi-text*; Isabelle et al. [1] refer to the set of techniques used to produce an aligned bi-text, or more generally to re-establish all or part of the translational equivalences between two texts, as *translation analysis*. In early work on translation analysis, the aligned segments of a bi-text were generally taken to be sentences, but in principle linkages between larger or smaller units are also possible; and, of course, one sentence in the source can be explicitly linked with two or more sentences in the target, and vice versa.

As for TransCheck, the basic idea is quite simple: as its name suggests, the system is intended to be used to check draft translations for certain kinds of translation errors, in much the same manner as a spelling checker is used to check monolingual texts for spelling errors. Once a translator has completed a draft translation, he or she submits the source and target language files to TransCheck for alignment; the system then proceeds to verify the aligned segments to ensure that they conform to certain basic and generally recognized properties of a good translation.

2. Source language interference

One such property is the absence of source language interference in the translation. Perhaps the most obvious and extreme instances of source language interference are presented by illicit borrowings, defined simply as the presence in a translated text of a SL word that is not recognized to be part of the TL lexicon. For languages that are as closely in contact as English and French are

1. Please see the acknowledgments at the end of this paper.
2. For details on the CITI's current alignment algorithm, which has a success rate of over 98%, see Simard et al. [7].

in Canada, illicit borrowings are a very common source of translation errors. Of course, all languages borrow constantly, and in some cases, the authorities are not in agreement as to whether or when a given foreign form has been naturalized. In most cases, however, there is no controversy. If words like “brunch”, “briefing”, or the verb “backer” are detected in a French text, the translator should be advised that these are unacceptable according to widely accepted standards of good written French. Existing French spelling checkers may be able to flag these forms as unknown words, provided they do not coincide with other legitimate French forms. What TransCheck does, in addition, is suggest to the user a correct French word that can be used in place of each such “anglicisme”.[3]

Deceptive cognates (or *faux amis*, as they are more colourfully called in French) constitute a slightly more subtle form of source language interference. These are etymologically related pairs of words in two languages that retain a degree of morphological similarity but which can no longer serve as mutual translations because they have come to have disjoint meanings. One classic example is the pair “library//librairie”: the French noun cannot be used to translate its English cognate because its meaning now corresponds to the English word “bookstore”. There are hundreds, if not thousands, of cases like these between English and French, many of them well documented in collections like [12], and we would like our translation checker to be able to automatically detect them in a draft translation. What exactly does this involve?

Recall that the first thing that happens when two language files are submitted to TransCheck is that the system automatically aligns them; that is, it explicitly links the segments in the two texts that are in a translation relation. What we now want TransCheck to do is to verify that none of these linked segments are composed of the pair “library//librairie”; and similarly for all the other deceptive cognates that have been entered into the system’s database of prohibited translations. At first glance, this would seem to require that the system generate correct word-level alignments. The problem is that even the best word alignment programs (e.g. Dagan et al. [14]) cannot ensure that every word in the source text will be linked to a corresponding word in the target. How can we be certain that the deceptive cognates we are searching for will not be among those words the program does not connect?[4] For this type of application, it seems preferable for a system like TransCheck to align segments at the sentence level, and to formulate queries that are slightly less precise: the system will flag any pair of aligned sentences in which the English segment **contains** the noun “library” or “libraries” linked to a French segment containing “librairie” or “librairies”. And while in principle a pair of aligned sentences could be flagged in which “librairie” was not actually the translation of “library” but of some other word in the English sentence, in practice it turns out that this kind of noise is very infrequent. In fact, the results of the preliminary tests we have conducted with TransCheck on more than half a million words of translated text seem to support the general viability of a translation checker that uses a **sentence alignment** program to detect prohibited **word correspondences**; see Macklovitch [6], section 3, for a detailed account of these results.

---

3. While in principle a translation checker may be bi-directional, TransCheck has a definite English-to-French bias, due principally to the fact that the reference works we consulted (see below) focus almost exclusively on problems of French usage.
4. In fact, it is not at all obvious how any word alignment program like that in [14] which is based on estimates of translation probabilities could learn to link pairs of words that are generally *prohibited* as translations.

In the preceding paragraph, we mentioned that TransCheck incorporates a database of prohibited translations, or what could be called an anti-lexicon. Not being experts on correct French usage ourselves, we decided to consult a number of well-known reference works on English-French translation problems, which are listed in the bibliography as [8-12]. In all, we extracted descriptions of over 2800 translation problems, corresponding to what these authorities maintain are attested and frequently occurring problems of source language interference in French translations. In addition to deceptive cognates and illicit borrowings, these include calques and certain target language improprieties. While the latter are not strictly speaking caused by SL interference, their formal description is identical to that of deceptive cognates: the only difference being that the pair of words involved do not derive from a common root. Thus "abreuvoir", according to one of our authorities, is incorrect when used in the sense of the English "drinking fountain." Of course, this same problem could arise in a French text that wasn't a translation. The advantage of having a source text at our disposal is that it provides TransCheck with a readily accessible indication of the sense of an ambiguous French word, in the form of the source word(s) it translates. As for calques, these are multi-word expressions that are translated literally from the source language, word for word, in a way that produces an unacceptable result in the target. Some, like "certificat de naissance", are perfectly correct in their syntactic form, as well as being semantically transparent, but simply do not correspond to the accepted or standardized term in the target language; in this case, "acte de naissance". Others, like "à la journée longue", calqued on the English "all day long", do contravene the rules of French grammar. The temporal adjective "long" does not usually follow the noun it modifies in French, and the correct translation of the English here is "à longueur de journée".

This, then, is how TransCheck operates to detect cases of source language interference in a draft translation. Each entry in the system's database of prohibited translations is converted into a query that is applied to the bi-text produced by the alignment program. Queries that are matched in the text are then flagged for the translator's attention during a TransCheck editing session. This will be easier to explain if we refer to the screen dump of TransCheck output on the following page. (We should mention, in passing, that the first TransCheck prototype is implemented on Sun workstations and uses Emacs as an interface, to allow programmers and linguists maximum flexibility). The top window shows the aligned bi-text in side-by-side format, with the current segment in bold and the forms matching the query underlined. The lower window contains the full database entry from which the compiled query has been derived, and the output of the grammatical analysis of the source and target sentences, which is currently limited to part-of-speech tagging. In this case, the type of error that has been flagged is a deceptive cognate: according to Van Roey [12], the French noun "déception" is not a permissible translation for the English noun "deception"; either "tromperie" or "duperie" should be used instead. Hence, in this instance, the aligned segment flagged by the system would seem to contain a true translation error: the French translation incorrectly renders the meaning of the English by something like "treachery, *disappointment*, lies...". If the translator concurred, he or she would modify the draft translation accordingly and then hit "n" to proceed on to the next potential error flagged by the system.

As mentioned above, the database of the first TransCheck prototype contains about 2800 entries like that shown for "deception//déception" in the lower window of the screen dump.

TransCheck : a screen dump

```
                         *TransCheck Text*
rs   File   Edit   Help
is all I ask.                          Je n'en demande pas plus.

u have a marketing and sales guy       Le représentant des ventes et de la
ll Dash-8s, where does he get          commercialisation des Dash-8, de qu
irection right now?                    reçoit-il ses directives en ce
                                       moment?

istory of the boundary, the A-B        L'histoire de l'établissement de la
and this body of water which is        frontière, de la ligne AB et de ce
eign and internal to Canada, is        plan d'eau, qui fait partie des eau
f a trail of bullying, broken          intérieures canadiennes, est marqué
ses, treachery, deception, lies        par l'intimidation, les promesses
reason, and it carries on to           brisées, la trahison, la déception
very day.                              ainsi que le mensonge, et ça se
                                       continue jusqu'à ce jour.

ll, the 1987 CPP changes               Les modifications apportées au RPC
ded the payment of two benefits        1987 permettaient, en outre, le
Emacs: *TransCheck Text*      (TransCheck)--66%-----------------------
le id=357>
r no=357.1>
piled query: f(déception+/NomC\*) e(deception+/NC\*)
nçais incorrect: déception+/NomC
lais:            deception+/NC
nçais correct:   tromperie ; duperie
e d'erreur:      faux ami complet
rce:             Van Roey
or>

>
eInfo: Couple ID = 357,
       Document = "text",
       Source = English
ens[L1] = {13567}
ens[L2] = {15913}
             DT                        L':                Déte-ddéf-fémi-s:
ry:          NC                        histoire:          NomC-fémi-sing
             PP                        de:                Prép
             DT                        l':                Déte-ddéf-masc-s:
ary:         NC                        établissement:     NomC-masc-sing
             PC                        de:                Prép
             DT                        la:                Déte-ddéf-fémi-s:
             ZZ                        frontière:         NomC-fémi-sing
             PC                        ,:                 Punc-pccm
Emacs: *TransCheck Info*      (TransCheck Narrow)--Top-----------------
ge collecting...done
```

Needless to say, this is nowhere close to being a complete inventory of all cases of source language interference between English and French. This first prototype has no pretensions of being a production system ready for field-testing, but is merely intended to illustrate how such a system might eventually operate. Moreover, we do not expect that eventual users will always agree with the opinions of the authorities whose works we consulted in compiling the database; in fact, it sometimes happens that the authorities do not agree among themselves. For this reason, it is absolutely essential that eventual end-users of TransCheck be able to extend and modify the contents of the system's database, so that it reflects their own standards and those of their clients. More generally, we see TransCheck as a tool that will help enforce translation norms – *in the plural*; because there surely is no single translation standard that applies across-the-board to all domains and types of text.

One final word on the current contents of the system's database. Of all the problems of SL interference described in references [8-12], we selected only those that could be characterized as absolute translation interdictions. Referring back to our earlier example, the English noun "library" can *never* be translated as "librairie" in French; whenever it is, this will constitute a translation error. However, the majority of attested translation errors described by our authorities do not exhibit this absolute character, but generally depend on various contextual factors: source word *x* can *sometimes* be translated as target word *y* if this or that condition elsewhere in the sentence is met; or inversely, *x* cannot be translated as *y* when this or that condition elsewhere in the sentence is met. Again, to illustrate with a concrete example, consider the verb "graduer" in French, which can translate the English verb "to graduate", but only in the sense of marking something with incremental measurements; the French verb cannot be used in the (normally intransitive) sense of completing a course of study. Now there are different ways in which this discrepancy between the two verbs could be described: either in terms of the transitive or intransitive distinction, or perhaps in terms of the permitted semantic features on the verbs' subjects and objects. But whichever way is chosen, the description will clearly involve more than just the verb's form and part-of-speech. As we have seen, however, this is all the grammatical analysis that TransCheck currently provides. Hence, in order for TransCheck to eventually be able to handle the many cases of partial (or contextual) translation interdictions, the system will have to be extended and augmented to include some sort of syntactic parser and perhaps a semantic analyser as well. These form part of the longer-term research goals of the TransCheck project.

3. Terminological Consistency

Another property of a good translation that is often cited as being desirable, especially for technical texts, is terminological consistency.[5] Lengthy or urgent technical texts are frequently divided up among several translators or freelancers; the reviser or translation co-ordinator is then left with the onus of merging these parts into a coherent whole. Ensuring terminological consistency is an important part of this job, and it can be quite onerous. The question is: Could this task be automated to some extent with the help of a tool like TransCheck?

---

5. Indeed, machine translation vendors commonly invoke terminological consistency as one important argument in favour of MT; but that is often because their systems are restricted to providing a single TL equivalent for every SL word, including the words of the common non-technical vocabulary that may require multiple equivalents.

Suppose we adopt an extremely naive definition of the notion of terminological consistency, namely, that each occurrence of a designated source term must be translated quite literally as the target term specified by the reviser. Suppose too that the reviser is able to enumerate the terms that he or she wants checked for consistency before beginning the revision of the draft translations; this could take the form of a text-specific glossary in which each entry is a simple term equivalence statement – nothing more than "e-term1 = f-term1." Once the various portions of the draft translation had been merged, a bi-text would be produced as before, by having TransCheck align the sentences of the source text with those of the target. The system would then convert the term equivalence statements in the reviser's glossary into a series of TransCheck queries, complete with their morphological variants,[6] and apply each in turn to the bi-text. Those aligned segments found to contain one of the specified source terms **but not** the corresponding target term given by the reviser in the glossary would be flagged by the system for later revision.

We have recently implemented this rather simplistic scheme and have just begun to explore some of its implications. The results of our investigation will be reported more fully in Macklovitch [13], but already it is quite clear that the main challenge for this approach will be maintaining a tolerably low noise level – where by "noise" here, I mean instances of the source term which are not rendered exactly as specified in the glossary, but which are nonetheless acceptable. Pronominalization of the specified target term offers an obvious case in point, as does the reduction of a multi-word term to its head, e.g. the truncating of "tireur d'élite" to "tireur" once the full term has been properly introduced in the discourse context. Both types of anaphora are quite common in technical texts, so that we could not consider exempting every TL segment that happened to contain a pronoun, say, without increasing the risk of leaving undetected the very cases of terminological inconsistency we are after. On the other hand, it is doubtful that we could ask the reviser to verify all the flagged segments that contain TL pronouns or truncated terms, as this would mean forfeiting too much of the time and effort TransCheck is intended to save.

Quite obviously, we are going to have to complicate our simplistic schema by introducing some sort of (partial) treatment of target language anaphora, which hopefully will allow us to reduce the noise level to tolerable proportions, while still helping the reviser ensure terminological consistency in highly technical texts.[7]

4. Other verifiable properties

Completeness is certainly a desirable property of all varieties of translation. No translator wants to discover, after his text has been sent out, that he has accidentally omitted a sentence, a paragraph, or even a whole page. (Alas, such things do happen.) Yet to the best of my knowledge, no software is available on the market today that could help a translator or reviser perform this kind of elementary verification. The only way such omissions can currently be detected is through a painstaking manual comparison of every sentence and paragraph in the two texts. A moment's

6. For this and for its part-of-speech tagging, the system incorporates monolingual dictionaries and rules of morphological analysis and expansion for both English and French.
7. If we insist on technical texts in this discussion, it is because this same property, which does appear essential, say, in a parts list, could well be considered undesirable repetition in expository or other non-technical varieties of text.

reflection should make it evident, however, that what the translator is actually doing in carrying out this verification, i.e. checking that every sentence in the source text has been provided with a translation in the target (before going on to verify the content of that translation) is essentially identical to what our alignment programs are intended to do. In principle, a correctly aligned bi-text should make such omissions apparent and draw the translator's attention to the fact that a given major segment in the source text corresponds to **no segment** in the target. Could a system like TransCheck not produce such bi-textual representations that would allow these omissions to be detected automatically?

Though conceptually it may be quite straightforward, the automatic detection of such omissions is not at all trivial to implement. The problem hinges on the robustness of the alignment algorithms. Upon encountering a region where the source and target texts don't "match" (in whatever technical sense the program defines), can the algorithm correctly determine whether it is dealing with an omission or whether it is dealing with an alignment that is not one-to-one? (And if it is dealing with an omission, can the program recover after the omitted segment to continue properly aligning the two texts?) The solution would seem to require alignment algorithms that incorporate some sort of fixed anchors between the source and target text – either cognates or (estimated) word translations. For this particular problem, robust word alignment programs like that of Dagan et al. [14] would appear to hold more promise than sentence alignment programs based uniquely on calculations of segment length.

Another extremely boring task for a translator or reviser is to verify that all numerical expressions in a source text have been correctly rendered in the target. Texts in domains such as economics or statistics can be packed full of such expressions, and the smallest error in one digit is tantamount to a serious mistranslation: not only is it extremely embarrassing for the translator, but it can undermine the credibility of the entire text. The reason why human revisers find this task so boring is that the "translation" of numerical expressions is generally so straightforward that it requires almost no intellectual effort; and yet every number must still be checked for the possibility of an error of transcription that does occasionally occur. Is this not exactly the kind of mechanical operation for which computers are better suited than humans? On the basis of the aligned sentences it has paired, a system like TransCheck should be able to verify that for every source segment containing a numerical expression, the corresponding target segment contains the equivalent numerical expression; and where it doesn't, that pair should be brought to the reviser's attention. In actual fact, the problem is not as trivial as I have suggested here[8]; nevertheless, we are convinced that even a rudimentary numerical component within TransCheck should be able to validate a large proportion of the numerical expressions in most texts, and are currently working to incorporate such a module into the system.

8. To illustrate just a few of the complications frequently encountered, there is the obvious problem of numerical expressions that are written in different forms or according to different standards, e.g. "seven o'clock" vs. "7 p.m." vs. "19 h"; which is why all such expressions will have to be normalized before being compared. Less obviously, the text in one language may use a numeral, e.g. the date "1994", where the translation properly refers to the same period by means of a non-numerical noun phrase like "last year".

5. Conclusion

In the preceding sections, we have described the current state of the first TransCheck prototype and outlined our plans for extending the capabilities of the system so that it can automatically validate an increasing number of generally recognized properties of a good translation. At the CITI, we view TransCheck both as a medium-term development project and as a long-term research project. Of course, our ultimate goal is to develop a complete model of translation, i.e. one that could account for all possible translational equivalences between a source text and its acceptable target language translation(s). For when we have that, our system will be able to detect all possible errors of translation: these would fall out naturally as the set of those translational equivalences that the model does not define as well-formed.

However, as Y. Bar-Hillel sought to demonstrate over 35 years ago, the development of such a translation model would require formalizing (or somehow allowing the machine to access) unlimited amounts of linguistic and extra-linguistic knowledge. That is why he believed that fully automatic high quality translation (or FAHQT, as he called it) would always remain an unattainable goal, at least for unrestricted text. Despite the undeniable progress that has been made in MT and NLP since Bar-Hillel first published his celebrated demonstration, few people today would contest the claim that the ultimate goal of a complete model of translation still remains elusively remote. From this, it follows that *all* machine and machine-aided translation systems will necessarily have to operate on a very partial understanding of the texts they handle for a long time to come. In this conclusion, I would like to elaborate on an argument first made in Isabelle [15] at an earlier MT Summit on the implications of this fact for two opposing approaches to the general problem of translation automation: the approach of systems that are oriented toward translation production (e.g. classic MT systems), versus systems (like TransCheck) that are oriented toward translation analysis.

Conventional MT systems seek to generate an acceptable TL version of a source text on the basis of their partial understanding of it. To what extent and under what conditions they succeed is an issue of some contention that we cannot begin to consider here. The one point we do want to stress, however, is that these production-oriented systems cannot choose to ignore the most difficult and least understood problems of NLP. MT systems do not generate partial translations; they generate translations that are partially correct. Like it or not, your MT system is going to have to produce *some* TL equivalent for every anaphor in the SL text, for example. The consequences of having to produce a complete TL translation on the basis of an incomplete understanding of the source are certainly well known to MT users: all too often, the raw outputs are incorrect (i.e. fail to convey the full meaning of the source text), or incoherent (i.e. fail to respect the grammar of the TL). And the cost of these failings is passed on to the human post-editor or, where no post-editing is done, directly to the unfortunate reader.

Systems that operate on the translation analysis model, on the other hand, do not seek to autonomously generate a TL text; rather, their goal is to help the translator-revisor become more productive by automating some part of the translation process, on the basis of a representation of translational correspondences that have been discovered linking the various segments of an existing translation to the segments of the source text. Now ideally, both types of systems should be based on a complete model of translation. (To fully revise a translation, quite clearly, requires no less knowledge than to produce a first draft: that is part of the reason why revisers are selected from among senior and not novice translators.) But unlike their MT cousins, systems that are oriented toward translation analysis *can* choose to ignore some of the more difficult NLP problems that are currently beyond our ken. Rather than resulting in garbled output, the cost here may simply be less output, or system silence on various translation problems that we may one day be able to tackle. Someone might decide, for example, to take the first TransCheck prototype and employ it solely to detect certain problems of SL interference in draft translations. Again, it is debatable whether and to what extent translators or revisers would actually find this useful today. But the fact that the translation model underlying TransCheck fails to incorporate a proper theory of anaphora will have little or no impact on the question. As Isabelle [15] forcefully argues, a major point in favour of the translation analysis approach is that it has allowed for the rapid development of promising support tools (like TransCheck) on the basis of relatively weak translation models.

As research progresses and our models of translation are gradually filled out and refined, TransCheck should be able to take over the automatic validation of more and more aspects of human translations; and in so doing, it will contribute to a reduction in the cost of producing high quality translation. The important point is that support tools like TransCheck do not have to await the perfection of formal translation theory to render useful service, but lend themselves readily to incremental improvement.[9] In this sense, TransCheck can be compared to monolingual spelling or grammar checkers that are also less than exhaustive. Having a text checked by such a system offers no guarantee that it is fully accurate and correct; on the other hand, whatever errors the system does manage to automatically detect will still contribute to improving the quality of the final text.

Acknowledgments

The TransCheck system described in this paper is the joint product of a number of people who work (or have worked) at the CITI, whose contributions I want to acknowledge here. They are Pierre Isabelle, Director of the CITI's Machine-aided Translation Group, Monique Chevalier, George Foster, Marie-Louise Hannan, Jean-Marc Jutras, François Perrault, Xiaobo Ren, Michel Simard and Caroline Viel. I express my gratitude to all these colleagues for the excellent work they have done. None is of course responsible for any errors I may have inadvertently introduced in this report.

9. Of course, there is nothing to prevent introducing the same insights and improvements into production-oriented systems. Our contention, however, is that their impact will be considerably less, because the wheat in such systems tends to be smothered by the chaff.

References

1. Isabelle, Pierre; Dymetman, Marc; Foster, George; Jutras, Jean-Marc; Macklovitch, Elliott; Perrault, François; Ren, Xiaobo; and Simard, Michel. *"Translation Analysis and Translation Automation,"* in **Proceedings of the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation**, Kyoto, 1993.

2. Gale, William; and Church, Kenneth. *"A Program for Aligning Sentences in Bilingual Corpora,"* in **Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics**, pp.177-184, Berkeley, California,1991.

3. Brown, P.; Lai, J.; and Mercer, R. *"Aligning sentences in parallel corpora,"* in **Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics**, pp.169-176, Berkeley, California,1991.

4. Harris, Brian. *"Bi-text: A New Concept in Translation Theory,"* in **Language Monthly**, no. 54,1988.

5. Isabelle, Pierre. *"Bi-text: Toward a new Generation of Support Tools for Translation and Terminology,"* CITI Technical Report, 29 pages, 1992. Also published in French in **Meta**, no. 37:4, pp.721-737, 1992.

6. Macklovitch, E. "*Using Bi-textual Alignment for Translation Validation: the TransCheck System ,"* in **Proceedings of the First Conference of the Association for Machine Translation in the Americas**, pp. 157-168, Columbia, Maryland, 1994.

7. Simard, Michel; Foster, George; and Isabelle, Pierre. *"Using Cognates to Align Sentences in Parallel Corpora,"* in **Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation**, pp. 67-81, Montreal, 1992.

8. Colpron, Gilles. **Dictionnaire des anglicismes**, Éditions Beauchemin Ltée, Laval (Qué.), Canada, 1982.

9. Dagenais, Gérard. **Dictionnaire des difficultés de la langue française au Canada**, Les Éditions Françaises Inc., Boucherville (Qué.), Canada, 1984.

10. De Villers, Marie-Éva. **Multidictionnaire des difficultés de la langue française**, Éditions Québec/ Amérique, Montreal, Canada, 1988.

11. Rey, Jean. **Dictionnaire sélectif et commenté des difficultés de la version anglaise**, Éditions Ophrys, Paris, 1984.

12. Van Roey, Jacques; Granger, Sylviane; and Swallow, Helen**. Dictionnaire des faux-amis français-anglais**, Duculot, Paris, 1988.

13. Macklovitch, E. "*Qu'est-ce que c'est au juste la cohérence terminologique?"*, to appear in the **Actes des Quatrièmes journées scientifiques du réseau thématique LTT**, Lyon, September 1995.

14. Dagan, Ido; Church, Kenneth; and Gale, William. *"Robust Bilingual Word Alignment for Machine Aided Translation,"* in **Proceedings of the Workshop on Very Large Corpora**, Columbus, Ohio, 1993.

15. Isabelle, Pierre. *"Machine-Aided Human Translation and the Paradigm Shift,"* in **Proceedings of the Fourth Machine Translation Summit**, Kobe, Japan, 1993.