



1. L'erreur : nature et évaluation

## The Assessment of Professional Translation Quality: Creating Credibility out of Chaos

Malcolm Williams

Volume 2, numéro 2, 2e semestre 1989  
L'erreur en traduction

URI : <https://id.erudit.org/iderudit/037044ar>  
DOI : <https://doi.org/10.7202/037044ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association canadienne de traductologie

ISSN

0835-8443 (imprimé)  
1708-2188 (numérique)

[Découvrir la revue](#)

Citer cet article

Williams, M. (1989). The Assessment of Professional Translation Quality: Creating Credibility out of Chaos. *TTR*, 2(2), 13–33.  
<https://doi.org/10.7202/037044ar>

# The Assessment of Professional Translation Quality: Creating Credibility out of Chaos

Malcolm Williams

## 1.0 Introduction

### 1.1 The Problem

Translators, translation companies and translation services of governments and international organizations must all be accountable for the quality of their product. Yet the question of how to judge that quality can never be answered, it seems, to everybody's satisfaction. How can one evaluate the quality of an industrial-scale translation program when, in the eyes of many academics and practitioners, translation quality assessment (TQA) is too subjective or too rigid to yield valid, reliable results? In the preface to his important work *Approaches to Translation*, Newmark makes his position on prospects for the development of credible TQA systems very clear: «Detailed schemes for assessing translation» are, along with certain other fields of translation research, «dead ducks — either too theoretical or too arbitrary».<sup>1</sup> And indeed, a number of evaluation systems designed over the years have been unable to meet the full range of professional requirements because

---

1. P. Newmark, *Approaches to Translation* (Toronto, Pergamon Press, 1982), p. x.

of the difficulties involved in ensuring that the evaluator is competent or objective enough to judge, trying to quantify quality or the lack of it, trying to reduce the length of time required to assess a translation properly, and applying evaluation criteria consistently to an intellectual product that is often of uneven quality and heterogeneous in form and content.<sup>2</sup> Some researchers remain optimistic nonetheless. While agreeing that TQA may never attain the degree of precision, objectivity and rigour required in the natural sciences, Wilss feels there is « no reason to argue in terms of translation-critical agnosticism or defeatism ».<sup>3</sup> More recently Larose, though stopping short of recommending precise standards, has entertained the possibility of creating an « indice de congruence » between source text and translation.<sup>4</sup>

## 1.2 The administrative requirements for TQA tools

Because of its mandate and the fact that every federal government department is accountable to the Parliament of Canada for the efficiency and effectiveness of its programs, the Canadian government's Translation Bureau has to espouse the optimism of Wilss and design and apply credible evaluation tools yielding cogent, highly professional assessments. Notwithstanding the difficulty of such an enterprise, the Bureau cannot be satisfied with an approach based solely on an overall assessment of the text and general statements of the type « the translation is accurate and reads well » or « the translator's rendering is awkward and imprecise ».

The Bureau's main role is to facilitate communication within the government and between the government and members of the public by providing translation services to Parliament and more than one hundred departments and agencies with offices in the national capital and across the country. The Bureau is thus a major player in efforts to pursue Canada's bilingualism policies and implement the spirit and letter of the Official Languages Act (1969, amended 1988), which provides for the equal status of English and French in the federal government and for the promotion of both languages in Canadian society at large.<sup>5</sup> In carrying out its mandate, the Bureau processes almost 300 million words of translation each year.

---

2. For a review of TQA schemes, see R. Larose, *Théories contemporaines de la traduction* (Sillery, Presses de l'Université du Québec, 1987), pp. 199-216.

3. W. Wilss, *The Science of Translation* (Tübingen, Gunter Narr Verlag, 1982), p. 226.

4. Larose, p. 289.

5. See Government of Canada, Department of the Secretary of State, *Annual Report 1988-1989* (Ottawa, Supply and Services Canada, 1989), pp. 81-85.

In order to ensure that an undertaking of these proportions is well organized and supported, a number of specialized services have been set up and developed over the years — among them terminology, training and, in particular, evaluation. The Evaluation Division has two main functions :

- to help managers staff positions, accredit freelance translators and assess their work, and process staff relations cases ; this involves the preparation and marking of examinations and the analysis of professional translators' work ;
- to provide senior management with information on the quality of the product the Bureau delivers ; this information is presented in the estimates submitted to Parliament each year as a key indicator of the effectiveness of the Bureau's activities.

In 1986, a team headed by the author of this article conducted a review of the Bureau's translation quality standards and assessment procedures, known as Sical (Système canadien d'appréciation de la qualité linguistique). Senior management subsequently approved the team's recommendations, and the third version of Sical came into effect.

The purpose of this article is to explain the thinking behind those standards and procedures, specifically in response to the criticism levelled against any systematizing of TQA and to the argument that Sical is appropriate for revision but not for evaluation.<sup>6</sup>

## **2.0 Evaluation criteria**

For a TQA system to resist such criticism, it must demonstrate reliability and validity.

### **2.1 Reliability**

It is reliable if the evaluator's decisions are consistent and if the evaluation criteria are stable. Verification of reliability would involve a search for defects in the measurement procedures themselves, defects which could lead to biases or undue variations in the findings over a period of time. For TQA purposes, are procedures in place to ensure that the evaluators do not fluctuate between excessive rigour and extreme flexibility ? Are quality requirements clearly enough defined for decisions on borderline cases to be made with consistency and ample justification ? Is the TQA expert always objective ?

---

6. Larose, p. 203 : « Cette deuxième version (du Sical) est orientée vers la révision-correction plutôt que vers l'évaluation-constatation. »

## **2.2 Validity**

Validity refers to the degree to which the findings permit inferences about the characteristics of the target population. TQA validity is the extent to which the translation samples evaluated are representative of the texts normally translated by the service and the degree to which the evaluator is then able to make judgments about the level of quality, the strong points and the weak points of the service on the basis of those samples.

These are the criteria that must be applied by the person designing translation quality standards and assessment procedures. Validity and reliability have to be maximized if the TQA system is to be accepted in the field, and it is against them that the various alternatives have to be measured.

## **3.0 Specific issues**

### **3.1 The customer's judgment : a reliable yardstick ?**

One of those alternatives, often advocated by those opposed to detailed assessment schemes, is customer acceptance. But does the customer—a manager in another department, for example—have sufficient competence in the two languages and the required consistency in assessing quality ? The customer may well base her evaluation on a reading of the translation alone—for readability, logical flow of ideas, clarity, format and idiomaticity—and perhaps on the reactions of her immediate colleagues. But the chances are that the accuracy of the translation will be gauged only over the long term. Legal problems caused by mistranslations and omissions, taxpayers complaining that their rights have been infringed because the quality of the translation is inferior to that of the original—adverse consequences such as these will probably not arise immediately. Furthermore, the customer's satisfaction with quality may well be coloured by her level of satisfaction with the turnaround time for delivery of the translation. In other words, quality of service can be confused with translation quality, and the customer's assessment will therefore not meet the basic requirements of reliability and validity. Translation quality standards applied by experienced translators are necessary if a coherent, comprehensive evaluation of translators or a translation service is being sought. This does not mean that the customer satisfaction survey should be discarded as an evaluation tool. On the contrary, it can help the translation evaluator to adjust his approach in light of customers' requirements.

### **3.2 What is an acceptable translation ?**

If we assume that a full-fledged TQA system is required in addition to customer approval, we then have to define minimal acceptability. But is this possible ?

Researchers have rightly insisted that the notion of acceptability in translation is relative. For example, Simpkin claims that there is no such thing, in any absolute or general sense, as a « good translation » :

That concept may have some validity in the pedagogic/academic context ; it has none whatever in the commercial and little, I suspect, in the literary. « Good » applied to a product means « appropriate » ; and that means « appropriate to a situation », which thus has to be defined or specified.<sup>7</sup>

Sager's view is similarly based on the notion of appropriateness :

There are no absolute standards of translation quality but only more or less appropriate translations for the purpose for which they are intended.<sup>8</sup>

The dilemma facing every evaluator is thus highlighted. Should he assess the quality of the translation *in vitro*, without reference to the customer's specific requirements concerning timeliness, language quality and accuracy, or should he weight his judgment in light of those needs and constraints ?

In response to this dilemma, we can perhaps look to management theory and systems analysis for solutions, and specifically to the distinction between closed systems (no external input) and open systems (interaction with the environment).<sup>9</sup> A closed TQA system would be one in which the evaluator bases his rating on the number of errors and on their gravity in absolute terms and in which there is a single standard of acceptability. An outright mistranslation is a major error, a loss of shade of meaning is a minor error ; an excessively Gallic formulation is a major error, a spelling error is a minor one, and so on. The attractiveness of such a system lies in its simplicity : the translation is good or bad, depending on whether a fixed number of major and minor errors has been exceeded. A more complex variation of this system allows for a certain number of points to be subtracted for each type of error.

However, the single-standard *in vitro* approach has drawbacks. First, it can lead the evaluator to focus on individual units of translation and neglect the fact that their sense and structure depend on a longer text and its context.<sup>10</sup> He may thereby place more value on a literal

- 
7. R. Simpkin, « Translation Specifications », *The Translator's Handbook* (London, ASLIB, 1983), p. 129.
  8. J. Sager, « Quality and Standards — The Evaluation of Translations », *The Translator's Handbook* (London, ASLIB, 1983), p. 121.
  9. F. Luthans, *Organizational Behavior* (Toronto, McGraw-Hill, 1985), pp. 81-82.
  10. Larose, p. 291 : « Chaque unité de texte n'a de sens que si elle est insérée dans une totalité textuelle. »

or semantic translation than on the communicative or functional type of translation advocated by such experts as Nida and House,<sup>11</sup> and make an erroneous judgment on the acceptability of the text as a result.

Second, how do we decide on the number of errors an « acceptable » translation may contain? What is the real difference in quality between the two-page translation containing 5 or 6 errors of various kinds and a translation of the same pages containing 11, or 12, or 13 errors?

The answer lies partly in Simpkin's statement on acceptability. Such a system is very useful for training, academic and examination purposes. In an examination, for example, it is important for candidates to be rated according to a single standard. The candidates know the cut-off point, (generally) accept the rules of the game and act accordingly.

What are required in the workplace are a range of acceptable standards and an « open system » approach to evaluation, one which does not function *in vitro* with no reference to the environment in which the translation was produced.<sup>12</sup> Yet, paradoxical though it may seem, such an approach must at the same time be based on rigorous quantified standards, since it is only through analysis of translation units and quantification of error that evaluations can be successfully defended.

In order to refine the quantification of error within the framework of a single standard, some systems allow for an overall assessment whereby strong points detected in style and the rendering of particularly difficult passages can justify adjustment, by up to 20 percent, of the initial result. The marking scheme of the Canadian Translators and Interpreters Council (CTIC) is a good illustration of this.<sup>13</sup> However, schemes of this kind are appropriate for examinations but cannot provide for effective pre-delivery or post-delivery quality control of

---

11. See Newmark, pp. 38-56, for a comparative study of these approaches; E. Nida and C. Taber, *The Theory and Practice of Translation* (Leiden: Brill, 1969); J. House, *A Model for Translation Quality Assessment* (Tübingen, Gunter Narr Verlag).

12. G. Kandler, « On the Problem of Quality in Translation » in E. Cary and R. Jumpelt, *La qualité en matière de traduction*, Actes du III<sup>e</sup> Congrès de la Fédération Internationale des Traducteurs (New York, Pergamon, 1963), p. 295: « Translations cannot be simply judged as right or wrong. We should rather have a scale of evaluation according to the degree of coincidence of the interpretability of the translation with the interpretability of the original and we should not forget that quality cannot possibly be assessed apart from the purpose of the translation. »

13. Canadian Translators and Interpreters Council, « Standard Examination Board: Marker's Guide » (1988), internal document.

translations as finished products. For no matter how many strong points a translation may have, the mere fact that it contains a significant error of meaning could cost the customer considerable financial loss and damage her organization's image. Three paragraphs of brilliant translation containing masterstrokes of formulation and interpretation cannot compensate for one critical error, any more than the exciting lines of a new car can make the product acceptable if the brakes do not work.

Further, such a scheme unfortunately risks exchanging the arbitrariness of a single quantified standard for the subjectivity of an overall evaluation of a translation. This would be welcomed by practitioners when ratings were revised upward but could well lead to conflict if the ratings were revised downward in light of «non-numerical» considerations.

### 3.3 Subjectivity and arbitrariness in language quality assessment

Related to the problem of the closed system is the premise that TQA is as subjective as literary criticism, where evaluation of the target language is concerned. Whereas a competent evaluator can make a watertight defence of most errors of meaning detected, he often finds himself on shaky ground in matters of target language assessment. There are a number of reasons for this : the loose structures of the spoken language are incorporated into written communications, changes in spoken language that have been adopted by the population at large have yet to make their way into authoritative dictionaries and grammars, and usage preferences as well as the value attached to writing skills vary according to region, educational level, social group and even organization. Experts are thus frequently embroiled in disputes over one another's purist or laxist stance. For example, there is still considerable reluctance to approve the use of « due to » as a prepositional phrase and of « hopefully » meaning « it is hoped that », even though they have become an integral part of modern English usage.<sup>14</sup> Other controversial cases are the dangling participle and the split infinitive. Although he may recoil in horror from these usages, the evaluator must bury his personal preferences and make a judgment, based on what he considers current acceptable usage to be and what he thinks is acceptable to the reader/customer, as to whether the expression or

---

14. See R. Copperud, *American Style and Usage : The Consensus* (Toronto, Van Nostrand Reinhold Co., 1980), pp. 115 and 185 ; S. Greenbaum and J. Whitcut, *Guide to English Usage* (London, Longman, 1988), pp. 227, 343-344 ; C. Millward and J. Flick, *Handbook for Writers* (Toronto, Holt, Reinhart and Winston, 1985), pp. 463 and 467.



word is a serious enough weakness to be called an error. As David Crystal advises, somewhat tongue in cheek,

...people who harangue the press about split infinitives are doing the language no service. They are, rather, promoting a spirit of uncertainty which will ultimately do far more harm. My view is that, if you have an obsession, keep it to yourself.<sup>15</sup>

The divergence of opinion on individual units of language underscores, once again, the care with which evaluators must use error analysis as a tool for assessment since judgments on individual points of contention may have little bearing on the quality needed by the user of the translation.

#### **4.0 The search for solutions : rigour and realism in quality standards**

The Translation Bureau requires a TQA system not just for academic, training and examination purposes but also for program effectiveness measurement and performance appraisal—a more flexible, open evaluation system in which the evaluator takes account of the environment surrounding the translation. Under such a scheme, the TQA expert would adjust the evaluation criteria as the circumstances influencing production of individual translations demanded.

#### **4.1 The main external factors at play**

##### *Purpose or end use*

Knowledge of the purpose of a translation will help the TQA expert decide what impact weaknesses in the translation might have and, consequently, how much revision is or was needed. Let us suppose, for example, that a translation was for information purposes only, for the benefit of one or just a few experts in a given field ; the potential value and cost of revision prior to delivery would have to be weighed carefully. The user is interested in getting a readable version of the source text which gives her the appropriate facts ; she does not want to wait too long for a perfect translation. In fact, a large number of minor errors in language and meaning scattered throughout the text will probably have no adverse impact at all, whereas delays in producing the document may render it useless.

---

15. D. Crystal, « A Case of the Split Infinitives », *English Today*, III (July-September 1985), p. 17.

There is therefore a need for a set of standards that includes a minimum acceptable quality level. By the same token, considerable investment in revision to achieve optimum quality is frequently cost-effective, as in the case of regulatory documents such as acts of Parliament and safety standards, user manuals, public information documents, and ministers' speeches, where mistranslations and unidiomatic expressions that might have been innocuous in a translation produced for internal information purposes could have significant material, financial or political consequences because of the intended use of such texts.

In short, it is the purpose of the translation that determines how much effort should be expended to ensure quality. If he is to make a realistic assessment, the evaluator must be able to determine that purpose with a reasonable degree of certainty.

### *Distribution*

How many people are going to read the translation and act upon the information contained therein? What level or quality of language is required, given the educational background and social environment of those people? The distribution factor is closely related to end use, and can have a significant bearing on quality requirements.

### *Customer*

The customer and the end user are not necessarily the same person. The quality objective in translating a memorandum for a Cabinet minister's signature may be higher than that aimed for in translating a similar document for the signature of a junior officer.

### *Deadline*

Every document is perfectible and can be reworked or improved if the time is available to do so. In the world of professional translation, however, deadlines are often short. The customer thinks of translation at the last minute, with press time only a few days or weeks away, and in some cases is prepared or is persuaded to accept unedited or unrevised work provided that the deadline is met. The customer herself then assumes responsibility for bringing the translation up to a publishable level of quality, even though it may not have been considered satisfactory in other circumstances. It is important for an evaluation scheme to take this possibility into account rather than simply providing for rejection of such products on the basis of high professional standards.

### *Source text*

Two other factors sometimes invoked by disgruntled translators are the poor quality and the level of difficulty of the source text. If the evaluator himself has difficulty with a source text, he may soften what might otherwise be a harsh assessment of the translator's efforts. In general, however, the source text is readable, and a competent professional translator should be able to handle even difficult material well.

### *End use and acceptability*

Purpose, distribution and the customer are closely interrelated factors and may be subsumed, for the sake of simplicity, under end use. End use is not unrelated to House's notions of text function and functional equivalence, but since it was not designed with the mass processing of translations in mind, her model does not explicitly provide for more than one high standard of acceptable quality.<sup>16</sup> For his part, Larose underlines the importance of incorporating in the TQA process a comparison of the author's and the translator's purposes so that the evaluator's assessment will be as objective as possible :

*L'évaluation doit s'effectuer en fonction de l'adéquation entre le but du traducteur et celui de l'auteur, et non entre le but du traducteur et celui qu'aurait choisi l'évaluateur lui-même.*<sup>17</sup>

In the context of pragmatic translation, the translator's purpose may often differ from that of the source text or that of its author : a scientific publication may be translated for the information of just a few specialists, and vice-versa. Therefore, it would perhaps be more accurate, in the federal government context, to state that translation adequacy must be gauged in light of the translation's end use, not on the basis of the choices the evaluator, with his own standards of excellence, may have made as translator or reviser.

End use is a key to evaluation in that it forces the evaluator to combine error detection and quantification with a question : « Given the purpose of the translation, what are the possible consequences of the weaknesses in this translation ? » or « In spite of these weaknesses, is the text as a whole acceptable given the customer's needs or specifications ? »

## **4.2 Consequence of error**

The concept of consequence of error, applied widely in occupational analysis and job classification, is seldom discussed by translation

---

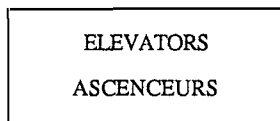
16. House, pp. 25-50.

17. Larose, p. 288.

experts, even those who, following Nida's example, favour an approach based on the effect of the translation on the reader.

Yet it is this concept which by and large determines the efficacy of a translation. The possible consequences of errors are numerous : they can affect national, organizational or personal safety and security ; they can cause financial loss ; they can cause embarrassment and loss of face. Such errors are costly and the evaluator's assessment will inevitably be a negative one.

Mistranslations in a non-essential part of a document (a footnote perhaps, or a bibliography) and minor language errors of various kinds (punctuation, spelling, usage, typing) are often insignificant and escape the reader's attention, even if there are a number of them. Yet the effect of these very same errors, occurring in an essential part of the text or in a translation read by hundreds of people every day, such as a sign or poster, can be devastating. For example, the misspelling in the following item—a sign in a Canadian government office complex—constituted a major error because of its visibility :



Clearly, the principle of equality of the two official languages had not been respected in this case.

### 4.3 Major and minor errors

In industrial quality control theory, errors have traditionally been broken down into three types by degree of gravity : critical, major and minor. An authoritative U.S. manual gives the following definitions :

**CRITICAL DEFECT.** A critical defect is a defect that judgment and experience indicate is likely to result in hazardous or unsafe conditions for individuals using, maintaining, or depending upon the product ; or a defect that judgment and experience indicate is likely to prevent performance of the tactical function of a major item such as a ship, aircraft, tank, missile or space vehicle.

**MAJOR DEFECT.** A major defect is a defect, other than critical, that is likely to result in failure, or reduce materially the usability of the unit of product for its intended purpose.

**MINOR DEFECT.** A minor defect is a defect that is not likely to reduce materially the usability of the unit of product for its intended purpose, or is a departure from established standards having little bearing on the effective use or operation of the unit.<sup>18</sup>

---

18. G. Hayes and H. Romig, *Modern Quality Control* (Encino (CA), Glencoe Publishing Co., 1982), p. 146.

« Generally speaking, » says Japanese expert Ishikawa, « one can never allow a critical defect, but a small number of minor defects is acceptable. »<sup>19</sup> The critical defect affects life and safety. A major defect prevents the proper functioning of a product, as in the case of a car engine that does not work. In other words, both major and critical defects have significant consequences. A minor defect such as a scratch on a car does not impair functioning but is not appreciated by the customer. By analogy, punctuation, spelling, typographical, grammar and usage errors and insignificant shifts in meaning may not be appreciated by the customer but in most instances do not jeopardize the usability of the translation.

On the basis of these considerations, a major or critical defect or error must have occurred in an essential part of a document for the evaluator to demonstrate convincingly that the usability of the translation has been reduced « materially ». So, for evaluations to be pertinent, an essential part of the text should be selected.

Critics have claimed that it is impossible to ascertain what is essential or unimportant to the customer. In fact, the TQA expert can judge the value and impact of the various parts of the source text and can, with the help of accompanying information such as the customer's request form, a work statement or written specifications, determine the end use of the text and the real needs to be met through the translation.

In order not to complicate its TQA system unnecessarily, the Translation Bureau has combined the critical and major defects into one category. However, the Bureau has found it practical to maintain the distinction between translation and language errors, while recognizing that an error of form can at the same time be an error of meaning and that a language error can cause a mistranslation or at the very least impede the reader's understanding of the translation. Given the above definitions, a major error in meaning, or major translation error, may be described as follows: *the complete failure to render the meaning of a word or group of words conveying an essential part of the message of the document*. Typical examples are a significant omission, a translation contradicting the meaning of the source text, or unintelligible jargon (assuming the source text was understandable).

A major language error is detected far less often, for reasons already stated—the evaluator's awareness of customers' varying sensitivities to language and the difficulty of imposing fixed rules on usage. It can be a rudimentary error (such as a spelling or morphosyntactical error on a public sign), an unwarranted neologism or the repetition of a rudimentary error (several spelling errors, or the same

---

19. K. Ishikawa, *What is Total Quality Control? The Japanese Way* (Englewood Cliffs (NJ), Prentice-Hall, Inc., 1985), p. 51.

punctuation error repeated several times in a 400-word passage) in an essential portion of the text.

Illustrations of the many types of major or minor translation and language error may be found in certain Translation Bureau documents<sup>20</sup> and in the model assessment in section 6. But the evaluator must always bear in mind that what may be a trivial slip of the pen in one set of circumstances may assume major proportions in another context and that an accumulation of minor errors in one sample is probably symptomatic of critical deficiencies elsewhere in the text.

#### 4.4 Sical

Many levels of acceptability could prove too cumbersome to manage effectively, while operating with one standard for all situations may force the evaluator to lower the acceptability threshold unduly in order to cover the many borderline cases that occur. Fortunately, the consequence of error criterion can provide parameters for a functional, multi-level acceptability grid.

The Translation Bureau has established three acceptable quality levels or ratings and one unacceptable level, which form the basis of Sical III, the Bureau's official TQA system. The narrative definition and numerical value (maximum number of errors tolerated) of each rating are presented below.<sup>21</sup>

##### *A rating : superior quality*

This level of quality would normally be required for texts in which even a small number of minor errors would not normally be acceptable because of the impact of the document and the consequence of error : e.g. publications, regulatory documents.

TRANSLATION/ LANGUAGE	The message of the source text is rendered accurately in correct, idiomatic language suited to the subject and end use.
TERMINOLOGY	Accurate and uniform.

---

20. See Government of Canada, Department of the Secretary of State, Translation Bureau, *Vade-mecum linguistique* (Ottawa, Supply and Services Canada, 1985); « Guidelines for Quality Controllers » (1986), internal document ; *Aide-mémoire d'autoperfectionnement à l'intention des traducteurs et des rédacteurs* (Ottawa, Supply and Services Canada, 1987).

21. For the official presentation of these definitions and numerical values, see Government of Canada, Department of the Secretary of State, Translation Bureau, «Management News» 9,4 (April 1, 1986), internal document.

*B rating : fully acceptable*

**TRANSLATION** No major error (complete failure to render the meaning of a word or passage that contains an essential element of the message ; also, mistranslation resulting in a contradiction of or significant departure from the meaning of an essential element of the message). May contain a few minor errors.

**LANGUAGE** No major error (incomprehensible, grossly incorrect language or rudimentary error in an essential part of the text).

General observance of the rules of language on which recognized authorities are in agreement. The translation reads well.

If the key requirements are for the translation to contain no errors of consequence and to read well, then we have described what constitutes acceptable quality for the customer in most situations. But we also know that errors considered insignificant in most contexts can barely be tolerated in certain high-profile texts, because of their impact. Readability and avoidance of significant errors in meaning cannot be the sole criteria for these translations, which must be of superior quality. The Translation Bureau has therefore set two acceptable quality standards for its finished product : « superior » and « fully acceptable ».

*C rating : revisable*

This standard represents the level of quality of a translation that could be made fully usable through cost-effective revision. This is the standard required of a translator whose work is regularly revised.

**TRANSLATION** The essential elements of the message are fairly well rendered, but the translation may contain one major error and/or several minor errors.

**LANGUAGE** The translation reads quite well, but the target language may depart from established norms and be unduly influenced by source language idioms, patterns and vocabulary.

**TERMINOLOGY** The translation may contain terms that have not been checked and therefore lack terminological accuracy or uniformity.

*D rating : unacceptable quality*

Cost-effective revision of a translation of this standard is impossible ; retranslation of many passages is necessary for the document to be fully usable.

### *Numerical value*

The quantitative grid is as follows :

Rating	Description	Quantified standard: Maximum number of major and minor errors per 400 words of source text
A	superior	Major            Minor 0                -                6
B	fully acceptable	0                -                12
C	revisable	1                -                18
D	unacceptable	>1              -              >18 (no maximum)

The quantitative standards are applied rigorously. At the same time, the TQA expert knows that it may be imprudent and unjust to give a translation a lower rating when the tolerance limit has been exceeded by only one or two errors and especially when language errors predominate. Consultation, review of the errors detected and further consideration of the external factors at play are recommended in such situations.

## **5.0 Procedures**

### **5.1 Teamwork**

The reliability of the Bureau's evaluation system is reinforced by the fact that the TQA experts take a team approach : evaluators work in pairs, so that an initial assessment is always reviewed by a second evaluator. Thus the risk of error is further reduced. Moreover, the TQA expert has the option of evaluating additional samples if, after analysis, he does not have enough clearcut evidence to make a confident judgment on translation quality.

### **5.2 Choice of passage**

It has been argued that sampling procedures are no substitute for revision of the whole translation prior to delivery. This brings up once again the thorny question of cost-effectiveness : the manager must determine whether the value added by comprehensive revision is equal to, greater than, or less than the cost of that revision, given the customer's needs and the translation service's own capacity.

The most accurate measure of length is the number of words translated, and the Bureau has determined that 400 words (about two pages of source text) is a manageable quantity on which to base its



evaluation procedures : it is long enough to present a significant, cohesive part of the document under assessment, and short enough to facilitate processing of texts on an industrial scale. It is important to select a sample containing an essential part of the message conveyed by the document and representative of the form and content of the whole ; accordingly pages containing lists, tables, graphs, bibliographies or extensive footnotes are, as a rule, avoided.

Yet 400-word sampling does not preclude the evaluator's examining other parts of the document (the introduction or appendixes, for example) and using the information thus obtained to analyse the samples.

Sical procedures thus provide an efficient means of assessing translations that are several thousand words long.

### **5.3 Selection of texts for evaluation of the Bureau's product**

We referred at the outset to the Translation Bureau's need for a mechanism by which to assess, as part of the program effectiveness measurement process, the quality of the product delivered to its clients. The mechanism devised should provide valid and reliable information on the quality of translations delivered by the Bureau.

Consequently the Continuing Evaluation System (CES) was developed. Each month translation services select translations according to a random numbers table (each translation request form bears a number). Each text selected is returned to the customer with an evaluation form for her to complete. At the same time, copies of the translation and the original are submitted to the Evaluation Division, as long as the material is suitable for evaluation, e.g. is not just a list of terms and is not a revision of a previous translation. Since a translation service can remove texts from the sample for such reasons, sampling procedures provide for services to submit more than the number of texts to be evaluated. Thus, in principle, all translations are to be evaluated by the customer as part of the Customer Satisfaction Survey (CSS), but a smaller number will be evaluated by the TQA team.

In this way, a statistically valid number of evaluations of the Bureau's product is built up during the year, and the resulting data will become an integral part of the Bureau's reports to Parliament.

### **5.4 Feedback**

The data generated by the CSS covers quality of service—meeting deadlines, communication with the customer, and translation quality as judged by the customer. The Bureau may use information from the survey to review its quality standards and TQA procedures, i.e. through a prudent comparison of results obtained from the CSS and the CES.

At the same time evaluations are returned to translation service managers for comment before ratings are finalized; since the evaluated party can contest ratings, the integrity of the overall evaluation process is enhanced.

## 6.0 Model sical assessment

### 6.1 Context

The French passage (the 400-word portion is marked off) is taken from a publication to be made available through Canadian embassies and consulates to potential tourists. A translation of superior quality is required because of the end use factor.

### 6.2 Source text and translation

The translation is coded as follows:

<i>Underlining (but no T or L)</i>	: weak point
<u>I</u>	: minor translation error
Ⓜ	: major translation error
<u>L</u>	: minor language error
Ⓛ	: major language error

Note: Strong points are indicated only on examination papers, where potential is being evaluated.

*La chasse à l'ours noir — 15 mai/15 juin*  
*Généralités*

L'ours quitte sa tanière, au terme de l'hiver, en assez bonne forme physique. Son premier souci est de rechercher de l'herbe verte pour se purger (exactement comme le fait un chien indisposé). Ceci peut l'amener à se déplacer assez loin en forêt primitive, uniformément composée de conifères.

À partir de là, ses réserves de graisse fondent rapidement parce qu'il ne trouve pratiquement rien à se mettre sous la dent: peu de végétaux, poissons difficilement accessibles, pas encore d'oiseaux nicheurs ou de jeunes des mammifères supérieurs de cette région. Sa principale ressource consiste en charognes, soit

*Black bear hunting — May 15 to June 15*  
*General:*

A bear leaves *its* (1.L) den at the end of winter in fairly good physical condition. *His* first concern is to find some green grass to purge himself with (exactly as a dog will do when it is unwell). This may lead him to travel some distance in virgin forest *comprised solely of* (2.L) conifers.

From then on his stored fat melts away (3.T) because he can find hardly anything to eat: plants are scarce, fish are difficult to catch and it is still too early for fledgling birds or the young of larger mammals in the area. His main food source consists of decaying meat, *either* (4.T + L) from the carcasses of ani-

en cadavres d'animaux morts de froid au cours de l'hiver, et notamment des ours malades, âgés ou dérangés dans leur tanière : ces charognes se décomposent lors du dégel, répandant une odeur que les ours flairent à des distances incroyables.

L'appât, mis en place en temps voulu, reproduit exactement ces conditions. En le réapprovisionnement régulièrement, on fixe un ou plusieurs ours dans sa périphérie.

À partir de mai, les ours trouvent naturellement d'autres ressources, mais ils continueront à visiter l'appât, tant par gourmandise que pour en interdire l'accès à d'autres ours en maraude. Ils y viennent donc deux ou trois fois par semaine, à partir de la tombée du jour.

Le rut de l'ours se situe en juin. Les mâles deviennent alors très mobiles, allant d'un appât à l'autre, tant pour trouver des femelles que pour écarter les autres mâles. Il en résulte des combats, avec les femelles, qui ne sont pas toutes réceptives ou sont encore accompagnées par leurs jeunes de l'année précédente (cas de cannibalisme occasionnels) dont les échos animent la quiétude des soirs de printemps en forêt !

À partir de cette époque on voit parfois des ours en plein jour à l'appât.

L'ours mâle dressé peut atteindre 1,80 m. Il pèse jusqu'à 180/210 kg, exceptionnellement 250. La femelle dépasse rarement 1,20 m et 90/110 kg.

La cible est donc apparemment importante, mais elle est en réalité étroite, ses contours n'étant que poil, crasse, graisse et chair non-vitale. Dans la mesure du possible,

mals that have frozen to death during the winter, and particularly bears that were sick, old or disturbed in their dens. The decaying carcasses decompose when there's a (5.T) thaw and give off an odour that bears can scent at incredible distances.

Bait put out at the proper time will give off exactly the same odour. If it is replenished regularly it will establish the boundaries (6.Ⓣ) for one or several (7.L) bears within range of its scent.

Beginning in May, bears naturally find other food sources, but they will continue to visit the bait, as much out of gluttony as to keep other marauding bears away. So they come two or three times a week, after sunset.

The rutting season for bears is in June. Males at that time become very mobile, going from one piece of bait to the other to look for females and chase away other (8.T) bears. Fights result, with the females, who are not all receptive or still have their last year's cubs with them (now and then there are cases of cannibalism), and their (9.) echos (10.L) liven the silence of spring evenings in the forest !

From then on bears are sometimes seen at the bait in broad daylight.

A dressed out (11.T) male can reach a size (12.T) of 1.80m and weigh from 180 to 210 kilos (13.L)—an exceptionally large one could (14.L) weigh 250. The female rarely exceeds 1.20m and 90-110 kilos.

The target therefore appears large, but is actually narrow, because its contours are only fur, layers of filth and fat, and non-vital flesh. Insofar as possible the hunter should aim

il faut le viser « en plein buffet », le plus près possible de la région de la gorge, l'idéal étant de lui casser la clavicule, parce qu'alors, même blessé, il ne peut aller loin (il tourne en rond !).

L'ours approche toujours furtivement de l'appât, se déplaçant sans bruit dans la forêt la plus dense. Son apparition est fugace : il saisit un morceau de l'appât et se retire le plus rapidement possible sous le couvert.

Les trappeurs indiens, actifs en forêt au moment où les ours prennent leurs quartiers d'hiver, savent mieux que quiconque, où fixer les ours au printemps. En outre les Indiens respectent l'ours, n'exagérant jamais le prélèvement sur un territoire. Ils affirment que l'ours sent la mort sur les territoires où ce prélèvement est excessif. Ils organisent aussi des postes d'affût extraordinairement astucieux parce qu'ils s'aident de la nature, alors que les Blancs veulent toujours la vaincre.

« dead centre », as close as possible to the throat area, the *ideal shot being to* (15.L) break the clavicle, because then (16.L) even when it is injured, it can't go far (it will go around in circles !).

A bear will always approach bait stealthily, moving noiselessly *through* the thickest part of the forest. Its appearance is fleeting ; it seizes a piece of bait and makes for cover as fast as its legs will carry it.

Indian trappers, who are engaged in forest activities when the bears go into hibernation, know better than anyone exactly where to find bears in the spring. Indians also have a respect for bears, never taking too many specimens in one particular region. They claim that bears smell death in regions where too many specimens have been taken. They also arrange extraordinarily clever observation posts because they let nature help them, whereas the white man is always trying to conquer nature.

### 6.3 Notes on errors and weak points

Refer to translation for the items concerned, which have been numbered accordingly.

- L 1. Switch from *its* to *his/him*.
- L 2. Correct form is *comprising* with no preposition.
- T 3. Omission : *rapidement*
- T + L 4. A serious error. Structure and meaning incorrect. *Either* must be followed by *or*, and *soit* here has the meaning of *namely* or *specifically*.
- T 5. Mistranslation : *du dégel* refers to *the* spring thaw.
- Ⓣ 6. Major translation error. The English reader can make little or no sense of this rendering at first. The French text, however, makes it quite clear that the odour will attract one or more bears within the vicinity or range of the bait.
- L 7. Gallicism : one or *more*.
- T 8. Imprecision : other *male* bears.

9. The antecedent of *their* is unclear — *fighters ? females ? cubs ? males ?* But the antecedent of *dont* in the source text is not clear either, so no error is assessed.
- L 10. Spelling error : *echoes*.
- T 11. Mistranslation. The measurements are those of a live bear. See *Webster's Third International Dictionary* (New York, 1981), 689, for senses of *dress out*. Borders on a major error.
- T 12. *Size* has a broader sense than the correct rendering, *length* or *height*.
- L 13. Inconsistency in use of metric symbols and abbreviations.
- L 14. Unnecessary change of auxiliary verb.
- L 15. Unidiomatic usage.
- L 16. Comma required.

NOTE : A number of weak points have not been assessed as errors and have been underlined but not commented on.

#### 6.4 Sical grid

ERROR COUNT				
	Ⓓ	Ⓕ	T	L
ERRORS:	1	0	6	9
TOTAL	1		15	
OUT OF 400	1		15	

RATING GRID (based on 400 words)

0	1-6	0	7-12	1	13-18	>1	>18
A		B		C		D	

RATING

C

#### 6.5 General remarks

Given the major error—the whole of the third paragraph is difficult to interpret—as well as a number of target language errors, the translation is of C-level quality : it must be revised to satisfy end use

requirements. Even if only a draft is required, the significant translation errors in the passage indicate that other major translation errors will probably be detected elsewhere in the document.

## **7.0 Conclusion**

The translation process is a complicated one, the individual product is often heterogeneous and of variable quality, customer requirements fluctuate, the environment changes with each text, and what seems a peccadillo one moment becomes an error of immense consequence the next. Against this backdrop of instability and complexity, the Translation Bureau has developed and implemented a reliable, valid TQA system by accepting a fundamental paradox: ratings based on error quantification are at one and the same time arbitrary and essential to TQA. TQA is not founded on consummate scientific rigour; it is an empirical system. With its vast experience in processing and evaluating thousands of translations and its knowledge of customer requirements, the Translation Bureau has defined the quality levels its services can reasonably be expected to attain on a day-to-day basis and has designed efficient TQA procedures that ensure uniformity and representativeness. Arbitrariness is further tempered as the evaluator exercises his own judgment and knowledge based not only on morphosyntactical considerations and « microtextual » analysis but also on the environmental factors affecting the translation process and the text as a whole. The result is a balance of rigour and flexibility which, when struck consistently by an experienced, competent team, resists criticism remarkably well.

Department of the Secretary of State of Canada