



## Les statistiques au service de la pédagogie de la traduction

Ginette Demers

---

Volume 5, Number 1, 1er semestre 1992

La pédagogie de la traduction : questions actuelles (1) et Miscellanées traductologiques

URI: <https://id.erudit.org/iderudit/037108ar>

DOI: <https://doi.org/10.7202/037108ar>

[See table of contents](#)

---

### Publisher(s)

Association canadienne de traductologie

### ISSN

0835-8443 (print)

1708-2188 (digital)

[Explore this journal](#)

---

### Cite this article

Demers, G. (1992). Les statistiques au service de la pédagogie de la traduction. *TTR*, 5(1), 101–111. <https://doi.org/10.7202/037108ar>

# Les statistiques au service de la pédagogie de la traduction

**Ginette Demers**

Il suffit de jeter un coup d'œil sur l'histoire de la traduction pour constater que, jusqu'au milieu du XX<sup>e</sup> siècle, c'est presque uniquement la traduction littéraire qui a intéressé les théoriciens. Leurs études ont souvent pris la forme d'une comparaison de textes traduits, mais l'analyse avait généralement comme objectif l'évaluation de la qualité du texte d'arrivée selon des critères le plus souvent fondés sur des impressions subjectives.

L'un des rares auteurs ayant manifesté un certain intérêt pour la traduction pragmatique est Schleiermacher. Dès 1813, en effet, «he differentiated between true translation, i.e. the rendering of artistic texts and texts having to do with the natural sciences, on the one hand, and mechanical translation, i.e. the transposition of pragmatic texts, on the other hand...» (Wilss, 1982, p. 31). Schleiermacher privilégie, cependant, le premier type de traduction et qualifie «d'interprétariat» — oral ou écrit — tous les cas où le traducteur n'est que «le simple serviteur d'un contenu objectif» (Berman, 1984, p. 233).

Depuis les années cinquante, toutefois, des facteurs comme la création d'organismes internationaux, le développement du commerce international, l'expansion des multinationales ont donné à la traduction pragmatique une importance sans précédent et nécessité, par conséquent, la formation d'un nombre croissant de traducteurs professionnels. On s'est vite rendu compte que la traduction non littéraire n'était pas aussi simple que Schleiermacher l'avait supposé, et les chercheurs ont

commencé à s'intéresser à ce genre de traduction. Contrairement aux théoriciens qui ont effectué des études sur la traduction littéraire, ils se sont rarement penchés sur l'observation du texte traduit. Ils ont plutôt tenté d'analyser le «processus de traduction» et ont élaboré, à partir de cette analyse, des théories destinées à servir de fondements à des méthodes visant à transcender les différences entre les types de textes. Il reste que ces méthodes sont fondées sur des hypothèses, et il arrive que les hypothèses soient contradictoires d'un auteur à l'autre. Vinay et Darbelnet postulent, par exemple, que «si nous connaissions mieux les méthodes qui gouvernent le passage d'une langue à l'autre, nous arriverions dans un nombre toujours plus grand de cas à des solutions uniques [... et que] nous pourrions même exprimer par un pourcentage le nombre de cas qui échapperaient encore à l'univocité» (1977, p. 24). Delisle estime, au contraire, que seules les formes et les expressions figées n'admettent qu'un équivalent et que «le contexte a pour effet de décupler les moyens linguistiques dont peut disposer le traducteur pour réexprimer en langue d'arrivée le sens du message original» (1980, p. 67).

Un autre exemple a trait à la traduction scientifique. Selon Juhel, en effet, «Si l'exégèse des messages est en général complexe, en revanche dans le discours scientifique, l'appréhension du sens est centrée sur un contenu informatif intrinsèque qui s'appuie sur un bagage cognitif précis. L'ambiguïté du texte scientifique, sa dimension émotive tendent vers zéro» (1982, p. 91). D'après Newmark, toutefois, «the common assumption that scientific and technical writing is concerned only with facts [...] is misplaced» (1969, p. 80). Cet auteur est d'ailleurs appuyé par House, qui ajoute «... seemingly "straightforward" ideational texts may contain subtle presuppositions and deliberate interpretations of facts...» (1977, p. 186).

De plus, certaines des hypothèses proposées par les théoriciens sont largement considérées comme des réalités, bien qu'elles n'aient jamais été vérifiées. Il suffit de penser à l'opinion très répandue selon laquelle le français, contrairement à l'anglais, tend à donner la prééminence au substantif. Or, comme le fait remarquer Delisle, «cette tendance n'a jamais été attestée objectivement par des méthodes statistiques rigoureuses» (1980, p. 177).

Ces quelques exemples montrent que l'enseignant de la traduction doit fonder sa pédagogie sur des bases qui sont souvent incertaines. Cette constatation m'a incitée à effectuer, à partir d'un corpus de textes pragmatiques et de traduction de ces textes, une analyse syntaxique et lexicale basée sur les statistiques. Comme ma spécialité est la traduction scientifique, je me suis surtout intéressée au domaine des sciences. Mon objectif était de voir, dans un premier temps, si la comparaison de textes scientifiques avec des textes appartenant à d'autres domaines permettrait de relever davantage de caractéristiques que ne le permet l'observation du seul domaine des sciences; de vérifier, dans un second temps, certaines hypothèses portant sur la traduction pragmatique en général et sur la traduction scientifique en particulier.

Cette analyse a fait l'objet d'une thèse de doctorat soutenue à l'Université Laval en mars 1989 et, comme la thèse est très volumineuse, je la remanie actuellement en un format se prêtant mieux à la publication. J'aimerais décrire brièvement cette étude, puis exposer les incidences des analyses de ce genre sur la pédagogie de la traduction.

## 1. Sélection du corpus

Le corpus observé était constitué de dix-sept extraits d'ouvrages pragmatiques anglais et de deux traductions françaises de chacun des échantillons. Sur le plan statistique, le fait d'observer deux traductions d'un même texte permettait de repérer les anomalies possibles et, par conséquent, d'obtenir des résultats plus précis que ne l'aurait permis l'observation de trente-quatre textes et de traductions uniques de chacun d'entre eux. Pour des raisons d'objectivité, les livres ayant fait l'objet de deux traductions ont été repérés dans l'*Index Translationum*, catalogue bibliographique publié par l'Unesco. L'*Index Translationum* a également été utilisé pour ce qui est de la répartition des ouvrages par domaine, et les extraits d'ouvrages figurant dans la rubrique «Sciences exactes, sciences appliquées» (6 livres) ont été comparés aux extraits de livres appartenant aux domaines «Histoire, biographie» (7 livres) et «Sciences sociales, pédagogie» (4 livres), aucun ouvrage appartenant aux autres domaines pragmatiques répertoriés dans l'*Index Translationum* n'ayant fait l'objet de deux traductions entre 1950 et 1980, étendue de la période d'observation.

J'ouvre ici une parenthèse pour faire remarquer que le nombre minime de livres ayant donné lieu à deux traductions (17 sur 12 415) et le fait qu'ils ont souvent été retraduits à la suite d'une mise à jour de l'original semblent indiquer que, dans les domaines pragmatiques, c'est plutôt par nécessité que par souci d'esthétique que les ouvrages font l'objet de traductions différentes. L'objectivité étant le critère principal de mon étude, je n'ai toutefois pas porté de jugement sur la qualité des traductions analysées. Il reste qu'il était peu probable que les traducteurs soient tous également mauvais. De plus, sur le plan statistique, les quelques maladresses, lourdeurs ou ambiguïtés apparaissant dans un corpus d'environ 15 000 mots n'ont pas d'influence sur les conclusions générales.

Parmi les différentes méthodes d'échantillonnage — échantillonnage aléatoire simple, échantillonnage stratifié, échantillonnage par grappes — celle qui se prêtait le mieux à un contexte linguistique était l'échantillonnage par grappes. J'ai donc choisi d'analyser les 300 premiers mots du troisième chapitre des livres retenus, chaque échantillon de 300 mots étant considéré comme une «grappe» du domaine étudié. L'inconvénient de cette méthode, c'est qu'elle présuppose que les échantillons sont représentatifs. Or, il est tout à fait possible que certains des extraits composant le corpus ne soient représentatifs ni du livre d'où ils sont tirés, ni du domaine auquel l'ouvrage appartient. C'est dire que les résultats enregistrés dans un ou deux extraits seulement ne sont valables que pour ces échantillons. La récurrence des résultats, par contre, surtout dans les cas où les mêmes phénomènes ont été observés par d'autres chercheurs, augmentent considérablement les probabilités, de sorte que des tendances peuvent être dégagées.

## **2. Méthode d'analyse**

Les points syntaxiques et lexicaux étudiés ont été choisis parce qu'ils servent de base à des observations ou à des assertions sur la traduction pragmatique en général, ou encore sur les caractéristiques du discours scientifique en anglais et en français.

Sur le plan de la syntaxe, les points suivants ont été retenus: nombre de phrases dans chaque extrait, longueur des phrases, genre de

phrase (simple ou complexe), type (déclarative, interrogative, impérative ou exclamative) et sous-type (phrase positive, phrase négative).

Les phrases ont ensuite été divisées en éléments fonctionnels, et l'analyse a porté sur le type et l'ordre de ces éléments, sur la personne du sujet, sur le mode, le temps et la voix des verbes ainsi que sur la fréquence des verbes pronominaux, des copules et des impersonnels. La fréquence d'emploi des relateurs logiques a également été étudiée.

Quant à l'analyse lexicale, elle a porté sur le nombre de mots dans chaque extrait, sur les similitudes entre le texte de départ et ses traductions, ainsi qu'entre les textes traduits; sur les différences, enfin, entre l'original et les traductions, puis entre les traductions elles-mêmes. Ces différences comprennent les ajouts, les omissions et les changements de catégorie grammaticale. Le taux de synonymie a également été relevé de même que la proportion de termes dans les extraits d'ouvrages spécialisés.

Les échantillons ont d'abord été analysés phrase par phrase. Les résultats ont ensuite été compilés extrait par extrait, puis domaine par domaine, et présentés sous forme de tableaux statistiques. (Deux cent quatre-vingts tableaux figurent dans la thèse.)

### 3. Résultats

Le recours aux statistiques a permis de dégager trois grandes tendances distinguant les extraits d'ouvrages scientifiques des autres extraits. La première de ces tendances est le *souci d'impersonnalité*, aussi bien dans les originaux que dans les traductions. Les moyens linguistiques qui servent à dépersonnaliser l'énoncé sont l'absence de sujets à la première et à la deuxième personne du singulier, l'extrême rareté de la deuxième personne du pluriel, l'emploi de la première personne du pluriel renvoyant à l'auteur, ou encore à l'auteur et au destinataire, la nette prédominance des sujets inanimés, la fréquence du passif dans les textes anglais et celle du pronom «on» dans les textes français, la fréquence des impersonnels et, enfin, le faible taux de possessifs.

Une deuxième tendance qui se manifeste clairement, en anglais comme dans les textes traduits, est le *souci d'éviter toute ambiguïté*. La cohérence et la clarté des textes sont d'abord assurées par des liens interphrastiques, les articulations logiques étant proportionnellement plus nombreuses que dans les extraits d'ouvrages non scientifiques. On remarque aussi la répétition des unités lexicales fortes, le très petit nombre de pronoms renvoyant à une phrase précédente ainsi que le faible taux de synonymie.

Le troisième fait saillant est la *constance des structures syntaxiques et lexicales*: l'écart entre les traductions et les originaux, de même qu'entre les textes français, est en général très faible, tant sur le plan de la syntaxe que sur celui du vocabulaire. Parmi les points donnant lieu à de grandes similitudes, on trouve le taux de phrases simples et de phrases complexes, la proportion de copules et d'impersonnels, le faible pourcentage de différences pour ce qui est des mots pleins, la fréquence d'emploi des noms, des verbes et des adjectifs spécialisés, la fréquence des éléments fonctionnels et des parties du discours, de même que l'ordre de présentation de l'information dans la phrase.

En somme, les seules variations importantes touchent les mots outils et elles sont souvent obligatoires, c'est-à-dire imposées par les différences inhérentes aux codes linguistiques en présence.

L'approche statistique a aussi permis de confirmer ou d'infirmer, chiffres à l'appui, des assertions touchant soit la traduction pragmatique en général, soit la traduction scientifique, soit le discours scientifique en anglais et en français. Il est intéressant de faire remarquer que, pour ce qui est des points étudiés, 50% des assertions non fondées sur l'observation d'un corpus sont infirmées. Entre autres, l'hypothèse selon laquelle une traduction est toujours plus longue que l'original (Vinay et Darbelnet, 1958, pp. 184-185; Juhel, 1982, p. 67) se trouve infirmée, la marge entre la longueur des traductions scientifiques et celle des originaux étant pratiquement nulle (0,5%) si l'on excepte la servitude que constitue l'expression de l'article zéro anglais par une forme linguistique en français. (Dans le cas contraire, l'écart s'élève à 5%.) La prétendue prédominance du substantif en français (Vinay et Darbelnet, 1958, pp. 84-87; Delisle, 1980, p. 177) est

également infirmée, la proportion de noms étant équivalente dans les deux langues, qu'il s'agisse des extraits d'ouvrages scientifiques ou des autres échantillons.

Par contre, les résultats de l'analyse correspondent dans une proportion de 85% aux résultats obtenus par des chercheurs ayant effectué, à partir de l'observation de corpus, des études sur le discours scientifique en anglais ou en français (Barber, 1962; Phal, 1968; Huddleston, 1971; Vigner, 1976; Johansson, 1978; Sager, 1980; Kocourek, 1982). Les caractéristiques du discours scientifique français se retrouvent donc, en très grande partie, dans les traductions analysées. C'est dire que, sans que soit exclu le fait qu'il puisse y avoir des différences entre un texte traduit et un texte rédigé spontanément dans la langue d'arrivée, l'étendue des divergences est beaucoup moins importante qu'on le suppose généralement.

La mise en regard a permis, quant à elle, de préciser des caractéristiques observées par les chercheurs qui se sont penchés sur le seul domaine des sciences et de dégager des traits communs à d'autres domaines. Ainsi, selon Kocourek (v. 1982, p. 52), les sujets à la troisième personne prédominent dans les textes scientifiques. C'est exact, mais ces sujets sont également prééminents dans les textes non scientifiques figurant dans mon corpus. Ce qui distingue les échantillons tirés de livres scientifiques, ce sont plutôt les particularités suivantes: 80% des sujets à la troisième personne sont des noms désignant des inanimés alors que, dans les autres extraits, 80% des noms sujets désignent des animés humains; dans les textes scientifiques, l'emploi du pronom «on» est quatre fois plus élevé et celui de «there», «it» et «il» impersonnels trois fois supérieur; enfin, le nombre de pronoms à la troisième personne renvoyant à une phrase précédente y est négligeable, tandis que, dans les autres échantillons, la proportion s'élève à 25%.

L'un des rares résultats n'ayant pu faire l'objet d'une comparaison, faute d'études statistiques à ce sujet, est la fréquence des parties du discours en français et en anglais. Il était donc possible que les similitudes observées soient attribuables au fait que les textes français étaient des traductions. J'ai donc effectué une seconde étude à partir de vingt textes scientifiques anglais appartenant à cinq domaines différents



et de vingt textes rédigés spontanément en français dans les mêmes domaines. Dans ce cas encore, la proportion de noms est équivalente dans les deux langues.

## Conclusion

J'ai à peine commencé à explorer un champ d'étude aux possibilités de recherches pratiquement illimitées. Pour ne citer que quelques exemples, la grille d'analyse que j'ai élaborée pourrait servir de fondement à des travaux permettant de dégager les similitudes syntaxiques et lexicales communes aux traductions de textes spécialisés ainsi que les caractéristiques propres à une spécialité en particulier; ou encore, les domaines spécialisés étant souvent très vastes, on pourrait utiliser la méthode pour établir les analogies et les divergences entre les sous-domaines; comme mon étude a montré qu'il existe des similitudes et des différences selon le degré de technicité des textes, il serait possible de mesurer l'ampleur des analogies et des divergences sur ce plan également.

À partir de ce qui précède, on peut facilement imaginer la répercussion des études statistiques comparatives sur la pédagogie de la traduction. Les caractéristiques stylistiques des divers types de traduction ayant été établies avec une exactitude mathématique, l'enseignant pourra appuyer sa pédagogie sur des bases solides et transmettre à l'étudiant, qui apprend encore trop souvent au hasard des textes, des techniques lui permettant d'arriver plus facilement à reformuler l'énoncé dans une langue idiomatique.

Déjà, certains des phénomènes observés dans ma propre étude m'ont permis, au cours des deux dernières années, de mieux structurer et de systématiser non seulement le cours de version scientifique mais encore un cours de version générale s'adressant aux débutants.

Les résultats de l'analyse montrent, par exemple, que les taux les plus élevés de différences obligatoires entre l'anglais et le français touchent les articles et les prépositions et que, dans le cas des substantifs, les divergences portent essentiellement sur les noms adjectivaux anglais. Or, ces trois points sont des sources d'erreurs très généralisées en première année. J'ai donc élaboré une série d'exercices sur la traduction de l'article zéro, sur l'emploi des prépositions en anglais et

en français ainsi que sur la traduction des noms adjectivaux. À la suite de l'exercice, l'étudiant traduit un texte choisi en fonction de la difficulté à l'étude de façon à pouvoir repérer le problème en situation réelle, c'est-à-dire dans un texte qui comporte aussi d'autres difficultés. Les étudiants éliminent leurs erreurs deux fois plus rapidement que lorsque les textes à traduire n'étaient pas accompagnés d'exercices.

Dans le cas du cours de version scientifique, les caractéristiques syntaxiques et lexicales communes à l'ensemble des extraits scientifiques sont exposées, les différences selon le degré de technicité des textes soulignées, de même que les points constituant des sources d'erreurs fréquentes, comme la répétition des unités lexicales fortes, l'étudiant ayant tendance à éviter la répétition; la proportion importante de phrases complexes dans les textes hautement spécialisés, l'étudiant ayant tendance à couper les phrases (peut-être pour éviter les erreurs de syntaxe); le taux de passif dans les textes anglais, proportionnellement plus élevé que dans les textes généraux, et qui donne souvent lieu à des erreurs de structure lorsque le passif est suivi d'un gérondif. La moyenne du groupe, un mois après le début du cours, est d'environ 10% plus élevée que la moyenne enregistrée au cours des quatre années précédentes.

Il est évidemment possible que ces résultats constituent une exception puisque ces méthodes ne sont appliquées que depuis deux ans. Il me faudra attendre encore quelques années avant d'avoir suffisamment de données pour prouver scientifiquement leur efficacité au moyen d'une étude statistique.

**Université Laval**

#### **Auteurs cités**

- BARBER, C.L. (1962). «Some Measurable Characteristics of Modern Scientific Prose», *Contributions to English Syntax and Philology*, Göteborg, Acta Universitatis Gothoburgensis, pp. 21-43.
- BERMAN, Antoine (1984). *L'Épreuve de l'étranger*, Paris, Gallimard, 311 p.

- DELISLE, Jean (1980). *L'Analyse du discours comme méthode de traduction*, Ottawa, Éditions de l'Université d'Ottawa, 282 p.
- DEMERS, Ginette (1989). *Constantes et variations en traduction*, thèse inédite, Québec, Université Laval, 1989, xxiii-576 p.
- HOUSE, Juliane (1977). *A Model for Translation Quality Assessment*, Tübingen, TBL Verlag Gunter Narr, 344 p.
- HUDDLESTON, Rodney D. (1971). *The Sentence in Written English. A Syntactic Study Based on an Analysis of Scientific Texts*, Cambridge, University Press, 344 p.
- JOHANSSON, Stig (1978). *Some Aspects of the Vocabulary of Learned and Scientific English*, Göteborg, Acta Universitatis Gothoburgensis, 57 p.
- JUHEL, Denis (1982). *Bilinguisme et traduction au Canada. Rôle sociolinguistique du traducteur*, Québec, Centre international de recherches sur le bilinguisme, 116 p.
- KOCOUREK, Rotislav (1982). *La langue française de la technique et de la science*, Wiesbaden, Brandstetter Verlag, 259 p.
- NEWMARK, Peter (1969). «Some Notes on Translation and Translators», *The Incorporated Linguist*, Vol. VIII, n° 4, pp. 79-85.
- PHAL, André (1968). «De la langue quotidienne à la langue des sciences et des techniques», *Le français dans le monde*, n° 61, pp. 7-11.
- SAGER, Juan Carlos et David Dungworth (1980). *English Special Languages*, Wiesbaden, Brandstetter Verlag, 368 p.
- VIGNER, Gérard (1976). «L'Initiation à l'expression écrite dans les langues de spécialité. L'objectivation», *Le français dans le monde*, n° 122, pp. 26-32, 41-42.

VINAY, J.P. et Jean Darbelnet (1977; 1<sup>ère</sup> éd.:1958). *Stylistique comparée du français et de l'anglais*, Montréal, Beauchemin, 331 p.

WILSS, Wolfram (1982). *The Science of Translation*, Tübingen, Gunter Narr Verlag, 292 p.