

La traduction automatique : un outil pour les traducteurs ?

Caroline Champsaur, Unité de Références et Terminologie, OCDE

RESUME

Ces dernières années, la traduction automatique a resurgi sous de nouvelles formes : projets de recherche, logiciels à destination du grand public et sites Web proposant des traductions en ligne immédiates et gratuites. Quel usage le traducteur professionnel peut-il faire de ces outils nouvellement mis à sa disposition ? Deux scénarios s'offrent à lui :

1. utiliser la traduction automatique en amont du processus même de traduction ;
2. intégrer la traduction automatique au processus de traduction en l'associant éventuellement aux mémoires de traduction.

Le premier scénario augmente de manière significative le temps de post-édition et peut lasser le traducteur. Le second risque de perturber le processus mental en jeu lors de la traduction, y ajoutant de nouvelles opérations mentales telles que la correction linguistique d'un énoncé « construit » par un logiciel ou la vérification de l'adéquation d'un énoncé cible avec un texte source. Il n'est pas aisé pour un traducteur de s'approprier un outil de traduction automatique. L'intégration de ce type d'outil présuppose que la traduction automatique fournisse un résultat de qualité satisfaisante ne nécessitant pas de multiples corrections et vérifications avec le texte source.

ABSTRACT

In recent years, machine translation has resurfaced in new forms: research projects, consumer software and websites providing instant and free online translations. How can a professional translator use these newly developed tools? Two scenarios are possible:

1. using machine translation before the translation process itself;
2. using machine translation within the translation process. It may also be associated with translation memory tools.

The first scenario significantly increases the time needed for post-edition and could rapidly bore the translator. The second one may disrupt the mental operations involved in the translation process, as the translator would have to correct grammatical errors in sentences "built" by the machine translation software and to ensure concordance between source and target languages while translating. Machine translation impacts on the translation process itself. It has to provide good quality results to be efficiently used.

MOTS-CLÉS

Traduction automatique, processus de traduction, outils d'aide à la traduction, utilisation, traduction automatique en ligne, traducteur.

KEYWORDS

Machine translation, translation process, CAT tools, use, online machine translation, translator.

Utiliser l'ordinateur pour dépasser la barrière des langues est un rêve ancien. Depuis la Deuxième Guerre mondiale, des sommes colossales ont été investies pour élaborer une « machine à traduire » qui aiderait les humains à se comprendre sans se soucier de leurs langues respectives. Ces dernières années, la traduction automatique a resurgi sur le devant

de la scène sous la forme de nouveaux projets de recherche mais aussi de produits commerciaux à destination du grand public et de services offerts sur Internet.

L'objet de cet article n'est pas de discuter la qualité des résultats obtenus par les logiciels de traduction automatique mais d'étudier de quelle manière les traducteurs s'approprient ces outils et en tirent un avantage.

1. La première palette des outils dans la sphère de la traduction

Les outils linguistiques touchant la sphère de la traduction se sont multipliés depuis l'utilisation massive des ordinateurs. Logiciels de traitement de texte, correcteurs orthographiques, dictionnaires électroniques sont venus depuis longtemps aider les rédacteurs et les traducteurs dans leur tâche. Différents outils se succèdent - dont ALPS, considéré par Kingscott (1999: 7) comme le « grand-père de tous les systèmes de mémoires de traduction » - et dans les années 80 apparaissent les deux logiciels d'aide à la traduction qui sont encore très largement utilisés de nos jours : Trados, créé en 1984 par Jochen Hummal et Iko Knyphausen en Allemagne, et Multitrans, créé plusieurs années plus tard par Gerry Gervais au Canada. D'autres outils analogues suivront. L'objectif de ces outils est d'assister les traducteurs lors du processus de traduction en retrouvant pour eux des passages préalablement traduits.

Chronologiquement, les logiciels de traduction automatique sont apparus avant les mémoires de traduction puisque les recherches dans ce domaine ont commencé dès la fin de la Deuxième Guerre mondiale. Réservée dans un premier temps au monde militaire, la traduction automatique devait garantir la compréhension (et éventuellement la rédaction) d'un texte rédigé dans une langue non maîtrisée par le locuteur. Ces logiciels étaient basés sur des règles linguistiques écrites par des linguistes qui analysaient la structure grammaticale de la phrase en langue source (arbre syntaxique), construisaient un arbre syntaxique équivalent en langue cible et remplaçaient les « feuilles » (mots) sources de l'arbre initial par des « feuilles » (mots) cibles. L'objectif des logiciels de traduction automatique était de donner un accès direct au contenu d'un message en langue tierce sans recours à un traducteur professionnel. Ils n'avaient pas pour but d'aider les traducteurs professionnels mais de s'en dispenser. À ce titre, ils ne faisaient pas partie à l'époque de la palette du traducteur.

Le rapport ALPAC (Automatic Language Processing Advisory Committee) (ALPAC 1966 ; Hutchins 1996), paru en 1966 aux États-Unis, estime que la traduction automatique n'est pas au point et qu'elle n'offre pas de perspectives. L'impact de ce rapport est décisif : les financements sont bloqués et les recherches dans le domaine de la traduction automatique suspendues.

2. La deuxième jeunesse de la traduction automatique : projets de recherches, logiciels grand public et sites en ligne

Ces dernières années, la traduction automatique a resurgi sous de nouvelles formes. Des projets de recherche ont vu le jour comme MOSES (Moses ; Koehn 2007) et PORTAGE (Conseil national de recherches Canada 2010 ; Foster *et al.* 2009). MOSES est un projet principalement financé par l'Union européenne lancé en 2005, alors que PORTAGE est un projet du Conseil national de recherches du Canada lancé en 2004. Ces projets reposent tous deux sur la « traduction automatique statistique » qui tire profit de la puissance de calcul des ordinateurs récents, de l'efficacité des algorithmes de « pattern matching » et d'apprentissage automatique et surtout de l'important volume de corpus bilingues disponibles électroniquement. En traduction automatique statistique, les corpus de texte sont analysés par le logiciel qui construit un « modèle statistique », la désambiguïsation lexicale étant effectuée à la fois par l'étude de la collocation au sein d'un texte monolingue et par la comparaison avec le texte cible aligné.

Le traducteur professionnel est à présent pris en compte comme utilisateur potentiel de ces projets. Ainsi, dans le projet PORTAGE, l'objectif affiché est de : « mettre au point une technologie qui permettra à un ordinateur de traduire d'une langue à une autre et d'appliquer cette technologie dans le but *d'accroître la productivité des traducteurs humains* » (Conseil national de recherches Canada 2010) (L'italique est de nous).

À côté de ces projets de recherche, différents logiciels à destination du grand public sont commercialisés depuis quelques années (ex. : Systran, Reverso de Softissimo, Personal Translator de Linguatrec, Traduction Express de @Prompt) offrant un accès facile et peu coûteux à la technologie de la traduction automatique pour tous, et notamment pour les traducteurs. La société SDL, quant à elle, propose un portail de traduction automatique (SDL BeGlobal) d'accès payant. Certains de ces systèmes sont basés sur des règles linguistiques, d'autres sont des systèmes statistiques, et d'autres encore des systèmes « hybrides », mêlant règles linguistiques et calculs statistiques.

Enfin, la vraie démocratisation en matière de traduction automatique est l'apparition sur Internet d'une multitude de sites proposant des traductions en ligne immédiates et gratuites : Systran, Reverso, Yahoo ! BabelFish, SDL FreeTranslation.com, ProMT-Online, Babylon, WordLingo, Translator (<http://translate.reference.com/>), Frengly.com, Lexicool, notamment. Parmi eux, le plus populaire est celui de Google : Google Translate. Depuis fin avril 2006, ce site, basé lui aussi sur une approche statistique, propose de la traduction automatique dans plus de soixante-quatre langues. Franz Och, chercheur chez Google Translate, estime que l'équivalent du contenu d'un million de livres est traduit chaque jour en

ligne sur ce site, ce qui ferait de Google Translate le plus important fournisseur de traductions de la planète (Och 2012)¹.

Pour des raisons de confidentialité, l'utilisation de ces services en ligne n'est pas recommandée dans un cadre professionnel. Cependant, ces systèmes viennent répondre à un besoin toujours croissant de compréhension d'informations en langues étrangères. En paraphrasant Stig Dagerman (Dagerman 1952), nous affirmons que *notre besoin de traduction est impossible à rassasier*.

Mais comment le traducteur professionnel tirera-t-il concrètement parti de ces logiciels de traduction automatique ? Lui sont-ils utiles ? Si oui, à quel moment du processus de traduction les utiliser ?

3. Comment la traduction automatique s'insère-t-elle dans le processus de traduction (humaine) ?

Si l'on envisage le processus de traduction lui-même comme « une enquête non linéaire » tel que décrit par Michel Rochard (Rochard 1999), le traducteur n'est pas un simple « passeur » d'une langue à l'autre, il doit pêle-mêle mettre en évidence les éléments logiques du texte original, comprendre la signification de l'énoncé en texte source, le déverbaliser, se référer à des informations extratextuelles (monde réel), choisir une formulation adéquate en langue cible ayant le même niveau de performance fonctionnelle que celle du texte source. Dans ces différentes démarches intellectuelles, comment utiliser la traduction automatique ?

Deux possibilités s'offrent au traducteur: utiliser la traduction automatique *avant* d'intervenir sur le texte à traduire ou bien *en complément* des autres outils d'aide à la traduction.

3.1 Premier scénario : Traduire automatiquement avant de traduire humainement

L'avantage d'utiliser la traduction automatique en amont du processus de traduction proprement dit est d'éviter de perturber les opérations mentales nécessaires pour traduire.

Si les résultats obtenus par la traduction automatique sont suffisamment satisfaisants pour être exploitables, on envisagera, de lui « sous-traiter » une partie du texte à traduire. L'objectif est alors de s'épargner une partie du travail en le faisant faire automatiquement par la machine et donc de gagner du temps. Le succès de cette démarche dépend grandement de la qualité obtenue par la traduction automatique et de l'utilisation finale du texte ainsi obtenu. Dans le cadre d'une utilisation officielle d'un texte en langue cible (et non pas d'un simple accès à la signification d'un énoncé source), il est possible, voire probable, que l'importance du travail de postédition nécessaire pour obtenir un résultat satisfaisant fera perdre

tout ou partie du gain de temps réalisé par le fait de ne pas avoir traduit certains passages.

Toutefois, le traducteur souhaitera éventuellement exploiter le « texte intermédiaire » obtenu en sortie du processus de traduction automatique pour le transformer en véritable traduction du texte source. Ce « texte intermédiaire » qui est le résultat de la traduction automatique est un texte utilisant des objets linguistiques de la langue cible sans en utiliser pleinement les règles d'utilisation. Les incohérences à corriger sont d'ordre grammatical (fautes d'accord), logique (rupture de sens), terminologique (choix d'un mot ou d'une expression cible inadaptée). Il s'agit aussi parfois d'omissions (passages non traduits) ou d'absences d'insertion d'éléments assurant la compréhension immédiate du texte par le lecteur. Ci-dessous quelques illustrations d'exemples effectués à l'OCDE avec Systran :

	Texte source	Texte traduit automatiquement	Traduction attendue
Incohérence grammaticale	The principles <u>agreed</u> by the Committee	les principes <u>a</u> <u>approuvé</u> par le comité	Les principes <u>approuvés</u> par le Comité
Incohérence logique	Guidelines cover situations <u>only</u> where supplies are bona fide	les directives couvrent des situations <u>seulement où</u> les approvisionnements sont véritables	les Principes directeurs couvrent <u>uniquement</u> des situations <u>dans lesquelles</u> les transactions sont effectuées de bonne foi
Incohérence terminologique	This problem should be <u>read</u> as...	Ce problème doit être <u>lu</u> comme...	Ce problème doit être <u>considéré</u> comme...
Absence d'élément de compréhension	<u>Attention</u> is drawn to...	<u>L'attention</u> est appelée sur...	<u>L'attention du lecteur</u> est attirée sur...

Tableau 1. Exemples de traductions effectuées avec Systran

Corriger ce type d'erreurs nécessite de retourner au texte source pour comprendre le sens du texte initial, le texte cible fourni par la traduction automatique n'étant généralement pas compréhensible sans référence au texte d'origine. Naturellement, plus le nombre et la variété des incohérences à corriger augmentent, plus le temps de postédition et l'énergie demandée au traducteur seront importants. Le traducteur risque de se lasser de ce type de corrections fastidieuses.

3.2 Second scénario : Utiliser la traduction automatique en complément

Pour éviter d'avoir à corriger de trop nombreuses erreurs de traduction automatique, le traducteur est libre d'en limiter l'utilisation et de s'en servir uniquement en complément des mémoires de traduction. Le scénario est alors le suivant :

1^{re} étape : Le texte à traduire est comparé aux mémoires de traduction => insertion de la version cible des passages déjà traduits.

2^e étape : Le texte à traduire restant est traduit automatiquement => conservation des résultats que le traducteur juge suffisamment corrects.

3^e étape : Le reste du texte, c'est-à-dire ce qui n'a pas été rendu par l'une ou l'autre technologie, est traduit par le traducteur professionnel.

À l'issue des deux premières étapes de ce scénario, le traducteur aura dans son texte des passages issus des mémoires de traduction mélangés à des passages traduits automatiquement. Il s'agit de deux types d'objets fort différents : le premier est le fruit d'un travail intellectuel humain ayant fait l'objet d'une validation, le second est une « construction » logicielle présentant éventuellement les incohérences listées plus haut. Le traducteur est obligé de différencier les deux types d'information et de les traiter différemment : les passages issus des mémoires de traduction sont a priori corrects syntaxiquement et, si les mémoires sont bien gérées, ils sont équivalents fonctionnellement² aux passages sources avec lesquels ils sont alignés, l'incohérence syntaxique des passages proposés s'ajoutant au risque d'inadaptation par rapport au texte source.

Parallèlement à ces opérations de vérification et de correction du texte proposé par les mémoires et la traduction automatique, le traducteur doit, naturellement, toujours effectuer les opérations mentales propres au processus de traduction : compréhension de l'énoncé source, mise en évidence de la logique interne au texte (déverbalisation), vérifications externes, formulation en langue cible. Dans la figure ci-dessous, ces opérations spécifiques sont représentées par les lettres L, D, F, T, S, celles inhérentes à l'utilisation de mémoires de traduction et de la traduction automatique étant représentées par les lettres V et C :

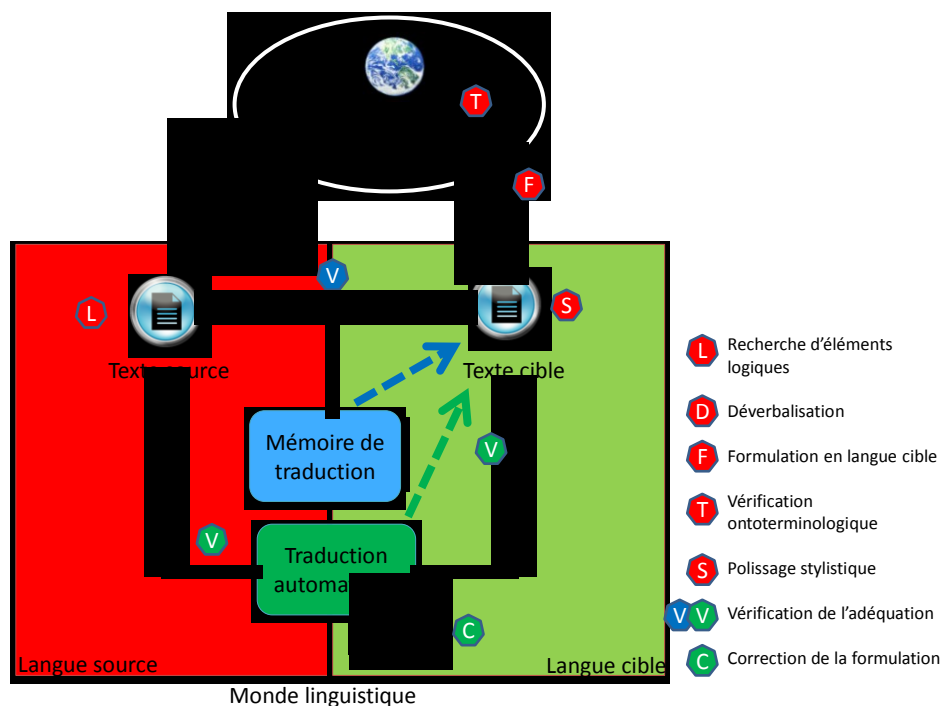


Figure 1. Opérations mentales en œuvre lors d'un processus de traduction appuyé par la traduction automatique et la mémoire de traduction

Devant l'augmentation de la variété et du nombre d'opérations mentales en jeu, il est légitime de se demander si les tâches de vérification et de correction des propositions de traduction ne risquent pas de se faire au détriment des opérations mentales propres à la traduction et si, à terme, le processus de traduction ne tend pas à se rapprocher à une démarche de postédition améliorée : correction d'erreurs, reformulation linguistique puis comblement des trous en traduisant les passages pour lesquels les systèmes automatiques n'ont fait aucune proposition. Ce risque, déjà présent avec l'utilisation des mémoires de traduction, est ici exacerbé d'une part à cause de la nécessité de corriger la syntaxe et l'orthographe des propositions de la traduction automatique et, d'autre part, à cause de la difficulté d'améliorer la qualité de ces propositions. Lorsqu'un logiciel de traduction automatique propose une séquence inadaptée, on corrigera immédiatement cette séquence dans la mémoire en l'amendant ou en la supprimant. Mais comme la traduction automatique repose sur des calculs statistiques, il faut introduire un très grand nombre de textes satisfaisants dans le corpus de référence pour peser sur la base statistique de calcul du logiciel. Les résultats obtenus par la traduction automatique sont donc difficilement améliorables par l'utilisateur. L'utilisation de systèmes de traduction hybrides (règles linguistiques + calculs statistiques) ne permet pas d'améliorer davantage les résultats obtenus puisque l'écriture de règles et leur utilisation sans effets de bord sont affaire de spécialistes.

Jugée peu prometteuse en 1966 par le rapport ALPAC, la traduction automatique renaît aujourd'hui de ses cendres et se propose même d'accroître la productivité des traducteurs. S'approprier cette technologie

n'est pourtant pas chose facile. Si les résultats proposés par la traduction automatique sont de qualité insuffisante, ils ne font pas gagner de temps dans le processus de traduction et risquent même de perturber la démarche intellectuelle mise en œuvre pour traduire. Malgré tout, il serait hâtif de rejeter la traduction automatique et de s'en désintéresser totalement. Les résultats obtenus par les grands projets de recherche MOSES et PORTAGE lancés de part et d'autre de l'Atlantique s'améliorent constamment et déboucheront sans doute sur un produit plus immédiatement utilisable. S'il parvient à l'intégrer à ses méthodes de travail, le traducteur se verrait alors offrir un nouvel outil prometteur.

Bibliographie

- **ALPAC** (1966). *Languages and Machines: Computers in Translation and Linguistics. A Report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council.* Washington, D.C.: National Academy of Sciences, National Research Council. http://www.nap.edu/openbook.php?record_id=9547&page=R1 (consulté le 16.12.12).
- **Anis, Jacques** (1994). « Repères bibliographiques en Traduction Automatique et Assistée par Ordinateur. » *Langages* 116, 123-124. http://www.persee.fr/web/revues/home/prescript/article/lgge_0458-726X_1994_num_28_116_1700 (consulté le 16.12.12).
- **Arnold Doug, Lorna Balkan, Siety Meijer, R. Lee Humphreys et Loisa Sadler** (1994). *Machine Translation: An Introductory Guide.* Manchester, UK : NEC Blackwell.
- **Champsaur, Caroline et Michel Rochard** (2011). « Ergonomie et appropriation : un exemple d'évolution de l'utilisation des outils de traduction ». *Traduire*, 224, 7-13.
- **Dagerman, Stig** (1952). *Vårt behov av tröst.* Norstedt & Söners, Stockholm. (1981). Traduction en français : *Notre besoin de consolation est impossible à rassasier* (tr. Philippe Bouquet). Le Paradou : Actes Sud.
- **Foster, George, Boxing Chen, Eric Joanis, Howard Johnson, Roland Kuhn, et Samuel Larkin** (2009). « PORTAGE in the NIST 2009 MT Evaluation. » *NRC-IIT Report.* <http://www.google.com/url?q=http://www.iro.umontreal.ca/~foster/papers/nist-techrep09.pdf&sa=U&ei=kjrPULmHGSwy0QXf7YDgDQ&ved=0CBQQFjAA&usq=AFQjCNHWI55oDWjfUR-MZsqUhxJJR8DNug> (consulté le 16.12.12).
- **Gaspari, Federico et John Hutchins** (2007). « Online and free! ten years of online machine translation: Origins, developments, current use and future prospects. » *Proceedings of the Machine Translation Summit XI.* Copenhagen (10-14 September), 199-206.
- **Hutchins, John** (1996). « ALPAC: the (in)famous report ». *MT News International* 14, 9-12. Reprinted in Sergei Nirenburg, Harold L. Somers and Yorick A. Wilks (eds) (2003). *Readings in machine translation*, Cambridge, Mass.: The MIT Press (2003), 131-135. <http://www.hutchinsweb.me.uk/ALPAC-1996.pdf> (consulté le 16.12.12).

- - (2005). « Current commercial machine translation systems and computer-based translation tools: system types and their uses. » *International Journal of Translation* 17, 1-2, 5-38. <http://www.hutchinsweb.me.uk/IJT-2005.pdf> (consulté le 16.12.12).
- **Hutchins, W. John** (1986). *Machine Translation: Past, Present, Future*. Chichester: Ellis Horwood.
- - **et Harold L. Somers** (1992). *An Introduction to Machine Translation*. London : Academic Press.
- **Kingscott, Geoffrey** (1999). « New strategic direction for Trados. » *International Journal for Language and Documentation*, November, 6-11.
- **Koehn, Philipp, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin et Evan Herbst.** (2007). « Moses: Open source toolkit for statistical machine translation. » *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*. Stroudsburg, PA: Association for Computer Linguistics, 177-180.
- **Och, Franz** (2012). « Breaking down the language barrier-six years in. » <http://googletranslate.blogspot.co.uk/2012/04/breaking-down-language-barriersix-years.html> (consulté le 16.12.12).
- **Rochard, Michel** (1999). « Traduction professionnelle et traduction pédagogique : le lien de l'enquête. » *Cahiers de l'Institut des langues et cultures de l'Europe (ILCE) Numéro 1: Pratiques professionnelles et pratiques d'enseignement*. Grenoble: Université Stendhal.
- - (2006). « La rétroconception entre compréhension et ré-expression des textes économiques. » Elisabeth Lavault-Olléon (éd.) *Traduction spécialisée : pratiques, théories, formations*. Bern/Frankfurt : Peter Lang.
- **Savoy, Jacques et Ljiljana Dolamic** (2009). « How effective is Google's translation service in Search? » *Communications of the ACM*, 52(10), 139-143. [doi:10.1145/1562764.1562799](https://doi.org/10.1145/1562764.1562799) (consulté le 16.12.12).
- **Slocum, Jonathan (éd.)** (1988). *Machine Translation Systems*. Cambridge: University Press.

Wébéographie

- **Moses.** <http://www.statmt.org/moses/> (consulté le 16.12.2012).
- **Conseil national de recherches Canada** (2010). <http://archive.nrc-cnrc.gc.ca/fra/projets/iti/apprentissage-machine.html> (consulté le 16.12.12).

Biographie

Caroline Champsaur est titulaire d'un DEA d'informatique fondamentale et est l'auteure d'une thèse de doctorat en linguistique théorique et formelle (Paris VII). Elle a participé au projet européen EUROLANG ayant pour objectif de développer un système de traduction assistée par ordinateur puis a encadré les développements de logiciels linguistiques dans une société d'édition. Depuis 1998, elle travaille au sein de la Division de la traduction de l'OCDE où elle dirige l'Unité de Références et de Terminologie.



caroline.champsaur@oecd.org

Notes

¹ « In a given day we translate roughly as much text as you'd find in 1 million books. To put it another way: what all the professional human translators in the world produce in a year, our system translates in roughly a single day. By this estimate, most of the translation on the planet is now done by Google Translate. » Franz Och, posté sur le blog <http://googletranslate.blogspot.fr> le 26 avril 2012.

² Pour une explication sur la notion d'équivalence fonctionnelle, voir Rochard (2006:104.)