

[Dossier]



Automatische vertaalsystemen: hel of hulp?

Joke Daems

Vertalers kunnen het huidige vertaalwerk niet bijhouden. Door de internationale economische groei en technologische vernieuwingen moet er steeds meer informatie vertaald worden naar een groter aantal verschillende talen. Die informatie moet ook nog eens snel beschikbaar gemaakt kunnen worden om relevant te zijn op het moment van publicatie. Om vertalers te helpen, wordt er dan ook steeds vaker gebruikgemaakt van automatische vertaalsystemen. Hoewel de output van zo'n systeem vaak te wensen overlaat, kunnen vertalers doorgaans sneller vertalen wanneer ze een automatisch vertaalde tekst verbeteren, ook wel postediten genoemd, dan wanneer ze al het werk zelf moeten doen. Maar sneller is niet noodzakelijk beter. Hoe zit het met de kwaliteit van zulke vertalingen? Is postediten niet vermoeiender dan manueel vertalen? En beleeft een vertaler nog wel plezier aan het vertalen bij het 'opkuisen' van automatisch vertaalde teksten?

Postediten en productiviteit

In de industrie hebben automatische vertaalsystemen hun nut reeds bewezen. Zulke systemen worden getraind op bestaande vertalingen van tekstsoorten die in een bedrijf regelmatig gebruikt worden. Het gaat dan vaak om technische teksten en handleidingen, teksttypes die doorgaans heel gelijkaardige formuleringen en woordenschat bevatten. Het softwarebedrijf Autodesk maakt al jaren gebruik van automatische vertaalsystemen waarvan de output vervolgens gepostedit wordt door hun vertalers. Zij rapporteren een stijging in productiviteit ten opzichte van gewoon vertalen van 30 tot wel 130 procent, afhankelijk van de talencombinatie (Plitt en Masselot, 2010; Zhechev, 2014).

Maar ook bij minder gecontroleerde teksttypes lijken vertalers baat te hebben bij het gebruik van automatische vertaalsystemen. In een aantal wetenschappelijke studies werd er gekeken naar het verschil tussen vertalers die algemene teksten, vaak krantenartikelen, moesten vertalen of postediten. Ook hier blijkt postediten een tijds winst op te leveren, al is die vaak minder uitgesproken dan bij het vertalen van technische teksten (Guerberof, 2009; Koehn, 2009; Daems, 2016).

Sneller en beter?

Wie al eens een tekst met Google Translate vertaald heeft, weet dat die vertalingen vaak heel wat te wensen overlaten. Zeker bij langere zinnen is de grammaticaliteit ver te zoeken en als een woord meerdere betekenissen heeft, slaagt Google er opvallend vaak in net de verkeerde betekenis te kiezen. Als de basis niet deugt, kunnen we dan wel verwachten dat een gepostedite tekst dezelfde kwaliteit heeft als een manueel vertaalde tekst? Onderzoek lijkt alvast uit te wijzen dat het antwoord op die vraag, misschien enigszins verwonderlijk, een volmondig 'ja' is. Er zijn zelfs studies waaruit blijkt dat de kwaliteit na postediting beter kan zijn dan die van

een gewone vertaling (Carl, Dragsted, Elming, et al., 2011; Garcia, 2011).

Bij postediten verwachten we dat de automatische vertaaloutput een impact heeft op de kwaliteit van de uiteindelijke vertaling en dus verwachten we vooral grammaticale fouten en betekenisproblemen in de gepostedite tekst. Uit de studie die de focus vormt van deze bijdrage (Daems, 2016), bleek echter dat er geen opvallende verschillen zijn in het aantal fouten bij postediten en manueel vertalen en ook het type fouten was niet statistisch verschillend bij beide vertaalmethodes. In deze studie moesten tien studenten uit de master Vertalen en dertien professionele vertalers enkele krantenartikelen vertalen en postediten. Betekenisproblemen ten opzichte van de brontekst vormden in beide gevallen de grootste foutencategorie. Opmerkelijk was ook het feit dat er geen grote verschillen leken te zijn tussen de professionele vertalers en de studenten, al had die laatste groep iets meer moeite met betekenisproblemen.

Is verbeteren vermoeiender dan vertalen?

Bedrijven zijn voornamelijk geïnteresseerd in het verhogen van de productiviteit en vertalers willen het liefst een product afleveren van hoogwaardige kwaliteit. Postediting lijkt op basis van onderzoek aan de wensen van beide partijen te voldoen, maar naast productiviteit en kwaliteit is ook inspanning een belangrijke factor. Als postediten immers vermoeiender blijkt te zijn dan manueel vertalen, houdt dat in dat een vertaler niet de hele dag kan blijven postediten, wat vermoedelijk de productiviteit noch de kwaliteit ten goede komt.

Een van de manieren waarop inspanning gemeten kan worden is door te kijken naar de oogbewegingen van een vertaler. Langere fixaties wijzen op grotere concentratie en dus meer mentale belasting. Intuïtief verwachten we langere fixaties bij het postediten,

omdat de vertaler moeite moet doen om naast de brontekst ook de output van het automatisch vertaalsysteem te begrijpen. Die intuïtie blijkt echter niet te kloppen. Fixaties zijn significant korter bij het postediten dan bij het manueel vertalen, wat zou kunnen wijzen op een mentaal minder belastende verwerking bij die eerste vertaalmethode.

De inspanning bij het postediten wordt wel beïnvloed door de fouten die aanwezig zijn in de automatische vertaaloutput. Zo heeft een vertaler meer tijd nodig om te postediten bij coherentieproblemen en stijgt de gemiddelde fixatieduur bij een groter aantal betekenisproblemen.

Wat denken de vertalers zelf? _ _ _ _ _

Op vlak van de objectieve metingen uit de voorgaande paragrafen lijkt postediting een interessant alternatief te zijn voor manueel vertalen, maar de subjectieve beleving van de vertaler speelt ook een rol. Als we immers vertrekken van het idee dat gelukkige werknemers betere werknemers zijn, is het belangrijk dat een vertaler plezier blijft beleven aan het vertalen in deze nieuwe vorm. Het klassieke idee is dat vertalers een hekel hebben aan postediten omdat ze hierbij gereduceerd worden tot schoonmakers van een slechte automatische vertaling en niet langer zelf creatief aan de slag kunnen. Maar klopt dit beeld vandaag nog steeds?

De deelnemers in bovenvermeld onderzoek hadden weinig tot geen ervaring met postediten. Voor en na het onderzoek werd er naar hun attitude ten opzichte van postediten gepeild. Er werden onderling heel wat verschillen vastgesteld. De voornaamste redenen die de vertalers voor deelname opgaven om geen gebruik te maken van automatische vertaalsystemen, waren dat ze er geen kennis van hadden en dat de output van zulke systemen een vertaler in een bepaalde richting duwt. Manueel vertalen was volgens sommige vertalers een creatiever proces; andere vertalers vonden het dan weer fijn dat er bij het postediten al iets stond om te verbeteren, wat de 'angst voor het lege blad' leek weg te nemen.

Opvallend was het feit dat de attitude van de vertalers, en dan voornamelijk die van de professionele vertalers, positiever was ten opzichte van postediting na afloop van het experiment. Een van de deelnemers vatte het als volgt samen:

'Voor dit experiment stond ik bijzonder sceptisch tegenover alle automatische vertalingen. Ik dacht dat het niet veel zou helpen, dat het me té veel in een bepaalde richting zou duwen, dat het meer een last dan een hulp zou zijn. Maar uiteindelijk viel dat beter mee dan ik had verwacht. Af en toe waren er handige, goed gevonden suggesties, en uiteindelijk deed ik nog altijd met de tekst wat ik zelf wilde.'

Hel of hulp? _ _ _ _ _

Wetenschappelijk onderzoek lijkt uit te wijzen dat postediting een valabel alternatief is voor manueel vertalen, ook voor algemene teksttypes als krantenartikelen, waar geen speciaal getrainde systemen voor bestaan. Postediten is doorgaans sneller dan manueel vertalen, de uiteindelijke tekst moet niet aan kwaliteit inboeten en het proces zou ook minder vermoeiend zijn dan manueel vertalen.

De automatische vertaaloutput wordt echter nog niet door alle vertalers als een hulp gezien en de voorkeur van de meeste vertalers ging na deelname nog steeds uit naar manueel vertalen. Toch leken vertalers na het postediten meer inzicht te krijgen in de mogelijkheden die automatische vertalingen hen kunnen bieden.

Een niet onbelangrijk aspect is ook het geld dat een vertaler verdient bij het postediten ten opzichte van het vertalen. In de vertaalmarkt is het helaas zo dat een klant doorgaans heel wat minder betaalt voor die eerste vorm van vertalen, wat vertalers mogelijk afschrikt. Er is meer onderzoek nodig naar de effectieve tijdswinst bij het postediten op langere termijn om zo tot correcte prijzen te kunnen komen voor deze vertaalvariant.

De systemen van de toekomst _ _ _ _ _

De ontwikkeling van vertaaltechnologie staat niet stil. Hoewel er in bovenvermeld onderzoek gekeken wordt naar 'statisch' postediten (vertrekken van een onveranderlijke automatisch vertaalde tekst) bieden huidige vertaalsystemen¹ ook interactieve en adaptieve automatische vertalingen aan. Deze systemen leren van de aanpassingen die de vertaler tijdens het postediten maakt, zodat latere zinnen beter vertaald worden en de vertaler op den duur ook minder vaak ergerlijke fouten moet verbeteren. Ook de 'statische' vertalingen worden steeds beter, of dat wordt toch beweerd. Zo werkt het nieuwe systeem van Google Translate met neurale vertalingen, wat volgens Google², afhankelijk van het taalpaar, vertalingen van 55% tot 85% betere kwaliteit zou opleveren ten opzichte van het oude systeem. Google rolde het nieuwe systeem voor het Nederlands³ uit tijdens het schrijven van dit artikel, zodat het meteen getest kon worden op de teksten uit bovenvermeld experiment.

Fragment uit brontekst:

The theory that high temperatures fuel aggressive and violent behavior is only just beginning to be studied. Using examples ranging widely from road rage, ancient wars and Major League Baseball, scientists have taken early steps to quantify the potential influence of climate warming on human conflict.

Oorspronkelijke GT-vertaling:

De theorie dat hoge temperaturen brandstof agressief en gewelddadig gedrag is slechts het begin te worden bestudeerd. De hand van voorbeelden op grote schaal, variërend van Road Rage, oude oorlogen en de Major League Baseball, hebben wetenschappers vroeg stappen om de mogelijke invloed van de klimaatopwarming op de menselijke conflicten kwantificeren genomen.

Nieuwe GT-vertaling:

De theorie dat hoge temperaturen agressief en gewelddadig gedrag branden, beginnen pas te worden bestudeerd. Met behulp van voorbeelden die veel ritten van woedingen, oude oorlogen en Major League Baseball, hebben wetenschappers vroegtijdige stappen genomen om de potentiële invloed van klimaatopwarming op menselijk conflict te kwantificeren.

Beter? Misschien, al heeft ook dit fragment duidelijk nog nood aan postediting. In de toekomst moet dus zeker onderzocht worden op welke manier deze nieuwe neurale vertalingen en de interactieve vertaalomgevingen het vertaal- en posteditingproces verder beïnvloeden.

Noten

- ¹ Een voorbeeld van een adaptief en interactief vertaalsysteem is Lilt: <https://lilt.com/>
- ² Zie blogpost <https://research.googleblog.com/2016/09/a-neural-network-for-machine.html>
- ³ <https://nederland.googleblog.com/2017/04/grote-verbetering-voor-het-nederlands.html?m=1>

Literatuur

- Carl, M., B. Dragsted, J., Elming, D. Hardt en A.L. Jakobsen (2011), The process of post-editing: A pilot study. In: B. Sharp, M. Zock, M. Carl en A. L. Jakobsen (red.), *Proceedings of the 8th International NLP/PCS Workshop*. Frederiksberg C: Samfundslitteratur, blz. 131-142.
- Daems, J. (2016), *A translation robot for each translator? A comparative study of manual translation and post-editing of machine translations: process, quality and translator attitude*. Gent: Universiteit Gent? Onuitgegeven proefschrift.
- Garcia, I. (2011), Translating by post-editing: is it the way forward? In: *Machine Translation*, 25, blz. 217-237.

Finge, R. (2017, 19 april). Grote verbetering voor het Nederlands in Google Translate [Google productblog] – <https://nederland.googleblog.com/2017/04/grote-verbetering-voor-het-nederlands.html?m=1>

Guerberof, A. (2009), *Productivity and quality in MT post-editing*. Paper presented at the MT Summit XII Beyond Translation Memories Workshop (WS3). Ottawa, Ontario, Canada.

Koehn, P. (2009), A Process Study of Computed Aided Translation. In: *Machine Translation*, 23, blz. 241-263.

Le, Q. & M. Schuster (2016, 27 september). A Neural Network for Machine Translation at Production Scale [Google Research Blog] – <https://research.googleblog.com/2016/09/a-neural-network-for-machine.html>

Plitt, M. en F. Masselot (2010), A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context. In: *The Prague Bulletin of Mathematical Linguistics*, 93, blz. 7-16.

Zhechev, V. (2014), Analysing the Post-Editing of Machine Translation at Autodesk. In S. O'Brien, L. W. Balling, M. Carl, M. Simard en L. Specia (red.), *Post-editing of Machine Translation: Processes and Applications* Newcastle upon Tyne, UK: Cambridge Scholars Publishing, blz. 2-23.

Joke Daems werkt als doctor-assistent voor het departement Vertalen, Tolken en Communicatie (faculteit Letteren en Wijsbegeerte) van de Universiteit Gent. Als onderzoeker is ze verbonden aan de onderzoeksgroep LT³ (Language and Translation Technology Team, lt3.ugent.be). Daarnaast biedt ze ondersteuning aan het Ghent Centre for Digital Humanities (www.ghentcdh.ugent.be) op vlak van digitale tekstanalyse.

e-mail: joke.daems@ugent.be