

## Tien pre-editeertips voor optimaal toegankelijke machinevertalingen



### Deel uw nieuws

Hebt u nieuws voor uw collega-taalprofessionals in Europa?

The Language Sector website maakt uw artikel, opinie of andere inhoud (geschreven in om het even welke Europese taal) zo toegankelijk mogelijk voor zo veel mogelijk taalprofessionals in zoveel mogelijk talen.

### Post-editing is goed

Met de optimale toegankelijkheid van uw informatie als doel maken we vanzelfsprekend gebruik van de technologieën die de taalsector zelf voortbrengt. Zoals u weet, verbetert de kwaliteit van de vertaalmachines snel. Maar waar nodig kunt u altijd elk artikel op The Language Sector website finetunen tot op het gewenste kwaliteitsniveau in de post-editingomgeving die in onze website is geïntegreerd.

### Pre-editing is nog beter

Met een goede post-editing kunt u de toegankelijkheid van uw informatie fors verbeteren. Dat klopt. Maar nog tijds- en kosteneffectiever dan een goede post-editing is een goede pre-editing van uw brontekst met het oog op een optimale vertaalbaarheid door de vertaalmachines. Met een beperkt aantal relatief kleine interventies kunt u de toegankelijkheid van de machinevertaalde informatie in uw bericht, artikel, opinie of andere inhoud zeer wezenlijk verbeteren.

### Top Tien Pre-Editing Tips

Hier zijn de Top Tien Pre-Editing Tips van onze huisexpert ter zake, Valentin Descamps. Valentin behaalde vorig jaar een masterdiploma in het vertalen aan de Universiteit van Mons in België. Hij is gespecialiseerd in de vertaling uit het Russisch en het Spaans in het Frans.

Voor The Language Sector website presenteert hij u tien eenvoudige maar effectieve technieken waarmee uw tekst vertaalmachinevriendelijker wordt én daardoor toegankelijker in vertaling.

Samengevat is zijn advies: **met 1 welgemikte pre-edit bespaart u de tijd en de kost van 20 post-edits** op The Language Sector website.

Niet aarzelen dus!





*Gent, 18 december 2020, Valentin Descamps* - Doel van het pre-editeren is een zo toegankelijk mogelijke machinevertaling te genereren in zoveel mogelijk verschillende talen. In de mate dat het nodig is, kan vervolgens de post-editor de machinevertaling “idiomatiseren”.

De tips die ik presenteer, verbeteren de machinevertalingen niet in alle talen in dezelfde mate. De voorbeelden in dit artikel laten zien dat de tips 9 op de 10 Franse en Spaanse vertalingen begrijpelijker maken, en 6 op de 10 Russische machinevertalingen. Het kan gebeuren dat een pre-editeertip die de Franse en Spaanse machinevertaling begrijpelijker maakt net het omgekeerde effect sorteert in de Russische vertaling.

Ik heb Google Translate gebruikt om het effect van de tips te evalueren. In de praktijk kunnen andere vertaalmachines een licht ander effect laten zien.

Als voor u de Russische vertaling om welke reden dan ook belangrijker is dan de Spaanse of de Franse, dan kunt u een pre-editing overwegen die meer specifiek op de Slavische talen focust.

Hierna presenteer ik mijn Top Tien Pre-Editing Tips. Elke tip illustreer ik met een paar voorbeelden. De voorbeelden hebben allemaal Nederlands als brontaal. Ze komen uit diverse en willekeurig geselecteerde teksten.

Elk voorbeeld toont in drie talencombinaties eerst de oorspronkelijke zin en vervolgens dezelfde zin met een pre-edit.

Wilt u graag reageren, hebt u een vraag of een suggestie? Neem dan gerust contact met mij op via [valentin.descamps@detaalsector.be](mailto:valentin.descamps@detaalsector.be). Ik kijk uit naar alle mogelijke feedback.



## Tip 1: vermijd afkortingen

Vermijd afkortingen. Afkortingen zijn voor de vertaalmachines vaak moeilijk vertaalbaar. Schrijf afkortingen voluit.

### Voorbeelden:

NL:

(1) “Het is belangrijk dat je die afk. kent.”

(2) “Het is belangrijk dat je die afkorting kent.”

NL → FR

« Il est important que vous le sachiez. »

« Il est important que vous connaissiez cette abréviation. »

NL → ES

“Es importante que lo sepas.”

“Es importante que conozca esa abreviatura.”

NL → RU

“Важно, чтобы вы это знали.”

“Важно, чтобы вы знали эту аббревиатуру.”



## Tip 2: vermijd zinnen zonder werkwoord

Zinnen zonder werkwoord zijn moeilijk voor de vertaalmachines. Voeg een werkwoord toe, anders riskeert er een werkwoord te ontbreken in de vertaling.

### Voorbeelden:

NL:

“Zo'n conversatietafel toch maar een beetje saai?”

“Is dit soort conversatietafels toch een beetje saai, volgens jou?”

NL → FR

« Une table de conversation comme celle-là un peu ennuyeuse, non ? »

« Ce genre de table de conversation est-il un peu ennuyeux, selon vous ? »

NL → ES

“Una mesa de conversación como esa un poco aburrida, ¿verdad?”

“¿Es este tipo de mesa de conversación un poco aburrida, según usted?”

NL → RU

“Такой стол для разговоров немного скучноват, правда?” (oké)

“ По-твоему, такой стол для разговоров немного скучен?” (Het eerste voorbeeld is beter)



### Tip 3: verduidelijk het geslacht van eigennamen

Personennamen kunnen specifiek zijn voor een taal of een land.

Voor een lezer zonder achtergrondkennis over het land en voor een vertaalmachine is het niet altijd duidelijk of een naam naar een man of een vrouw verwijst.

Tip: gebruik in de onmiddellijke omgeving van eigennamen zo vaak mogelijk mannelijke of vrouwelijke verwijzende woorden, zodat de machine het gender van de persoon mogelijk uit de context kan afleiden.

#### Voorbeelden:

NL:

“Anouk werd maandagmiddag in Brussel opgepakt door de politie.”

“Anouk, een jongedame van 19, werd maandagmiddag opgepakt.”

NL → FR

« Anouk a été arrêté par la police à Bruxelles lundi après-midi. »

« Anouk, une jeune fille de 19 ans, a été arrêtée par la police → Bruxelles lundi après-midi »

NL → ES

“Anouk fue arrestado por la policía en Bruselas el lunes por la tarde.”

“Anouk, una joven de 19 años, fue arrestada por la policía en Bruselas el lunes por la tarde.”

NL → RU

“Анук был арестован полицией Брюсселя в понедельник днем.”

“Анук, 19-летняя девушка, была арестована полицией Брюсселя в понедельник днем.”



#### Tip 4: vermijd te lange zinnen

Hoewel neurale vertaalmachines veel langere zinnen kunnen verwerken dan de vorige generatie (op statistiek gebaseerde) machines, is het toch beter om echt lange zinnen op te splitsen.

#### Voorbeelden:

NL:

“In de toekomst wil ik graag nog meer focussen op bedrijven, omdat ik er echt in geloof dat zij er enorm veel baat bij kunnen hebben, als ze de realiteit van spreektaal op de werkvloer erkennen.”

“In de toekomst wil ik graag nog meer focussen op bedrijven. Ik ben ervan overtuigd dat zij er enorm veel baat bij kunnen hebben, als ze de realiteit van spreektaal op de werkvloer erkennen.”

NL → FR

« À l'avenir, j'aimerais me concentrer encore plus sur les entreprises, car je crois vraiment qu'elles peuvent énormément bénéficier si elles reconnaissent la réalité du langage familier sur le lieu de travail. »

« À l'avenir, je voudrais me concentrer encore plus sur les entreprises. Je suis convaincu qu'ils peuvent énormément bénéficier s'ils reconnaissent la réalité de la langue parlée au travail. »

NL → ES

“En el futuro, me gustaría centrarme aún más en las empresas, porque realmente creo que pueden beneficiarse enormemente si reconocen la realidad del coloquialismo en el lugar de trabajo.” (oké)

“En el futuro, me gustaría centrarme aún más en las empresas. Estoy convencido de que pueden beneficiarse enormemente si reconocen la realidad de la lengua hablada en el lugar de trabajo.”

NL → RU

“В будущем я хотел бы еще больше сосредоточиться на компаниях, потому что я действительно верю, что они могут получить огромную выгоду, если признают реальность разговорной речи на рабочем месте.”



“В будущем я хотел бы еще больше сосредоточиться на компаниях. Я убежден, что они могут получить огромную пользу, если они признают реальность разговорной речи на рабочем месте.” (Het eerste voorbeeld is beter)

### **Tip 5: Speel met de structuur**

Hou de structuur van je zinnen zo helder mogelijk. Probeer geen twee ideeën in één zin te proppen. Als twee zinnen een logisch verband met elkaar hebben, gebruik dan verbindingswoorden om de relatie aan te geven.

#### **Voorbeelden:**

NL:

“Onze tarieven per woord variëren van € 0.10 tot € 0.16, doch voor de meeste projecten geldt het tarief van € 0.12 voor opdrachten van meer dan 2000 woorden.”

“Onze tarieven per woord variëren van € 0.10 tot € 0.16, maar voor de meeste projecten, die bestanden van meer dan 2000 woorden bevatten, geldt het tarief van € 0.12.”

NL → FR

« Nos tarifs par mot varient de 0,10 € à 0,16 €, mais pour la plupart des projets, le tarif de 0,12 € s'applique pour les travaux de plus de 2000 mots. »

« Nos tarifs au mot vont de 0,10 € à 0,16 €, mais pour la plupart des projets, qui contiennent des fichiers de plus de 2000 mots, le tarif sera de 0,12 €. »

NL → ES

“Nuestras tarifas por palabra varían de 0,10 € a 0,16 €, pero para la mayoría de los proyectos se aplica la tarifa de 0,12 € para trabajos de más de 2000 palabras.”

“Nuestras tarifas por palabra oscilan entre 0,10 € y 0,16 €, pero para la mayoría de los proyectos, que contienen archivos de más de 2000 palabras, la tarifa será de 0,12 €.”

NL → RU

“Наши ставки за слово варьируются от 0,10 евро до 0,16 евро, но для большинства проектов ставка 0,12 евро применяется к заданиям объемом более 2000 слов.”

“Наши ставки за слово варьируются от 0,10 до 0,16 евро, но для большинства проектов, которые содержат файлы, содержащие более 2000 слов, ставка будет составлять 0,12 евро.” (Het eerste voorbeeld is beter)





## Tip 6: speel met de context

Moderne vertaalmachines bekijken ieder woord in de context van (minstens) de hele zin. Introduceer je een moeilijk woord, introduceer het dan in een context die veel gelieerde informatie bevat die de vertaalmachines gemakkelijk kunnen verstaan. Zo verhoogt u de kans dat ze de zin correct zullen vertalen.

### Voorbeelden:

NL:

“Wie gevorderd wordt, heeft theoretisch niet te kiezen”

“Wie door de politie opgeroepen wordt, kan theoretisch niet weigeren te komen tolken.”

“Wie door de politie opgeroepen wordt, kan theoretisch niet weigeren in beide talen te komen tolken.”

NL → FR

« En théorie, celui qui est avancé n'a pas le choix »

« En théorie, quiconque est appelé par la Police ne peut refuser de venir interpréter. »

NL → ES

“En teoría, quien sea avanzado no tiene otra opción.”

“En teoría, cualquiera que sea llamado por la Policía no puede negarse a venir a interpretar.”

NL → RU

“У продвинутого нет теоретического выбора”

“Те, кого реквизировала полиция, теоретически не могут отказаться прийти и интерпретировать.” (niet oké)

“Те, кого вызывает полиция, теоретически не могут отказаться приходить и переводить на оба языка.” (oké)”



## Tip 7: speel met synoniemen

Als een heel specifiek woord echt moeilijk blijkt voor de vertaalmachines, stel u dan een beetje soepel op en wees bereid om een synoniem te overwegen. Denk aan de optimale **toegankelijkheid** van uw informatie in zo veel mogelijk talen.

Is dat ene super precieze woord met die ene super precieze betekenis echt zo belangrijk dat het gerechtvaardigd is dat u er twintig vertaalproblemen mee creëert en twintig post-edits noodzakelijk maakt?

### Voorbeelden:

NL:

"Dankzij deze vertraging en als we dit goed opvolgen, is het mogelijk om onder 2.000 intensieve zorgbedden te blijven de komende weken. Dit kan de ziekenhuizen ook wat tijd kopen om zich verder te organiseren en de snelle toestroom aan patiënten op te vangen", legde Van Gucht nog uit."

(2) "«[...]», verklaarde Van Gucht nog."

(3) "«[...]», verklaarde Van Gucht eveneens."

NL → FR

(1) «"[...]», a expliqué Van Gucht toujours absent. » (niet oké)

(2) « "[...]", a déclaré Van Gucht toujours. » (niet oké)

(3) « "[...]", a déclaré Van Gucht aussi. »

NL → ES

(1) "«[...] », explicó. Van Gucht todavía está fuera." (niet oké)

(2) "«[...] », dijo. Van Gucht todavía. (niet oké)

(3) "«[...] », dijo Van Gucht también."

NL → RU

(1) "«[...] », Ван Гухт объяснил." (niet volledig vertaald, maar oké)

(2) "«[...] », Ван Гухт объяснил."



## Tip 8: Vermijd taalspecifiek idioom

In het Nederlands sturen de mensen hun kat wanneer ze niet op een afspraak verschijnen. De meeste vertaalmachines vertalen zulk idioom dan letterlijk ("zijn kat sturen"), en dat is in vertaling vaak niet begrijpelijk voor de lezer. Het is beter in concrete, expliciete taal te schrijven.

### Voorbeelden:

NL:

(1) "Hij heeft zijn kat gestuurd."

(2) "Hij is niet naar de afspraak gekomen."

NL → FR

(1) « Il a envoyé son chat. » (nonsens)

(2) « Il n'est pas venu au rendez-vous »

NL → ES

(1) "Envió a su gato". (nonsens)

(2) "No acudió a la cita"

NL → RU

(1) "Он прислал свою кошку." (nonsens)

(2) "Он не пришел на приём."



## Tip 9: corrigeer taalfouten

Leg de lat hoog. Wees streng voor de kwaliteit van je eigen tekst. Taalfouten in de brontekst kunnen tot taalfouten in de doelttekst leiden. Dus, check nog eens extra je grammatica, spelling, interpunctie en de samenhang van je brontekst. Er zijn tools die jou daarbij kunnen helpen.

### Voorbeelden:

NL:

(1) “Goe kennen spellen en schryven is nog steeds belangrijk, zo heeft recentelijk onderzoek aangetoont. Op een sollicitatiebrief me spelfouten wordt je keihard afgerekent.”

(2) “Goed kunnen spellen en schrijven is nog steeds belangrijk, zo heeft recent onderzoek aangetoond. Op een sollicitatiebrief met spelfouten word je hard afgerekend.”

NL → FR

(1) « Comment savoir épeler et écrire est toujours important, selon des recherches récentes. Vous serez facturé très dur sur une lettre de motivation avec des fautes d'orthographe. »

(2) « Être capable de bien épeler et écrire est toujours important, selon des recherches récentes. Vous serez jugé sévèrement sur une lettre de motivation avec des fautes d'orthographe »

NL → ES

(1) “Cómo saber deletrear y escribir sigue siendo importante, según ha demostrado una investigación reciente. En una carta de presentación con errores ortográficos se le cobrará muy fuerte.”

(2) “Ser capaz de deletrear y escribir bien sigue siendo importante, según ha demostrado una investigación reciente. Serás juzgado con dureza en una carta de presentación con errores ortográficos.”

NL → RU

(1) “Недавние исследования показали, что умение писать и писать по-прежнему важно. За сопроводительное письмо с орфографическими ошибками с вас будет взиматься очень большая плата.”

(2) “Как показали недавние исследования, умение правильно писать и писать по-прежнему важно. Вас будут жестоко судить по сопроводительному письму с орфографическими ошибками.”



## Tip 10: schrijf in één taal

Is je tekst in het Nederlands? Schrijf zo veel mogelijk Nederlands. Dat klinkt logisch, maar sommige mensen gebruiken veel Engelse woorden en uitdrukkingen in hun Nederlandse teksten. Geen enkel probleem voor de Engelse vertaalmachine, maar wel vaak des te meer voor de andere machines, die dan minder goed presteren (in het Frans en in de andere Romaanse talen bijvoorbeeld).

### Voorbeelden:

NL:

- (1) "Dat was echt goed out-of-the-box gedacht van die Jan of whatever zijn naam was."
- (2) "Dat was een onorthodox goed idee van die Jan of hoe hij ook weer heette."
- (3) "Dat was echt een heel origineel gedacht van die Jan, of hoe hij ook weer heette."

NL → FR

- (1) « C'était vraiment bien de penser à Jan ou quel que soit son nom. » (compleet verkeerd)
- (2) « C'était une bonne idée peu orthodoxe de Jan ou quel que soit son nom. »
- (3) « C'était vraiment une pensée très originale de Jan, ou quel que soit son nom. » (Oké)

NL → ES

- (1) "Eso fue realmente bueno pensar en Jan o como se llame." (compleet verkeerd)
- (2) "Esa fue una buena idea poco ortodoxa de Jan o como se llame."
- (3) "Ese fue realmente un pensamiento muy original de Jan, o como se llame." (Oké)

NL → RU

- (1) " Это было действительно хорошее нестандартное мышление о Яне или как там его зовут." (verkeerd)
- (2) "Это была неортодоксальная хорошая идея от Яна или как там его звали."
- (3) "Это действительно была очень оригинальная мысль Яна, или как там его звали." (Oké)

