

The use of parallel corpora in linguistics

Annemarie Verkerk

Translation: Online and offline, losses and gains

Nijmegen, June 25-26 2012

Parallel corpus

a collection of texts that are all translations of a single original text that is made accessible in some way

Parallel text

The boy's name was Santiago. Dusk was falling as the boy arrived with his herd at an abandoned church. The roof had fallen in long ago, and an enormous sycamore had grown on the spot where the sacristy had once stood.

He decided to spend the night there. He saw to it that all the sheep entered through the ruined gate, and then laid some planks across it to prevent the flock from wandering away during the night. There were no wolves in the region, but once an animal had strayed during the night, and the boy had had to spend the entire next day searching for it.

He swept the floor with his jacket and lay down, using the book he had just finished reading as a pillow. He told himself that he would have to start reading thicker books: they lasted longer, and made more comfortable pillows.

It was still dark when he awoke, and, looking up, he could see the stars through the half-destroyed roof.

I wanted to sleep a little longer, he thought. He had had the same dream that night as a week ago, and once again he had awakened before it ended.

He arose and, taking up his crook, began to awaken the sheep that still slept. He had noticed that, as soon

Erster Teil

I

Der Jüngling hieß Santiago. Es fing bereits an zu dämmern, als er mit seiner Schafherde zu einer alten, verlassenem Kirche kam. Das Dach war schon vor geraumer Zeit eingestürzt, und ein riesiger Feigenbaum wuchs an jener Stelle, an der sich einst die Sakristei befand.

Er entschloß sich, die Nacht hier zu verbringen. Er trieb alle Schafe durch die zerfallene Tür und legte einige Bretter so davor, daß ihm die Tiere während der Nacht nicht entweichen konnten. Zwar gab es keine Wölfe in jener Gegend, aber einmal war ihm eines der Tiere während der Nacht entkommen, und er hatte den ganzen folgenden Tag mit der Suche nach dem verirrtten Schaf verbracht.

Dann breitete er seinen Mantel auf dem Fußboden aus, legte sich nieder und nahm das Buch, in dem er gerade gelesen hatte, als Kopfkissen. Vor dem Einschlafen überlegte er sich, daß er in Zukunft dickere Bücher lesen sollte, weil man länger etwas davon hatte und weil sie eine bequemere Kopfstütze abgaben.

Es war noch finster, als er erwachte. Als er nach oben schaute, sah er die Sterne zwischen den Dachbalken durchscheinen.

›Eigentlich wollte ich noch weiterschlafen‹, dachte er bei sich. Wieder hatte er den gleichen Traum gehabt wie

ParaSol parallel corpus

<p>78458 All three of you will receive detentions -- yes , you too , Mr Longbottom , nothing gives you the right to walk around school at night , especially these days , it ' s very dangerous — and fifty points will be taken from Gryffindor . "</p>	<p>Այո՛ , պարոն Լոնգբոթ՛ թո՛մ , դուք նույնպես : Ոչ մի բան ձեզ իրավունք չի տալիս պարետային ժամի արգելանքը խախտելով , գիշերով թափառել ամրոցում , հատկապես այս օրերին , երբ իրավիճակն այնքան վտանգավոր է դարձել շրջակայքում ... Եվ հիսուն միավոր կհանվի Գրիֆինդորի հաշվից :</p>	<p>Sie alle werden Strafarbeiten bekommen - ja , auch Sie , Mr Longbottom , nichts gibt Ihnen das Recht , nachts in der Schule umherzustromern , besonders dieser Tage ist es gefährlich - und fünfzig Punkte Abzug für Gryffindor . " " Fünfzig ? "</p>
<p>78842 Slytherins , on the other hand , clapped as he walked past them , whistling and cheering , " Thanks Potter , we owe you one ! "</p>	<p>Իսկ սլիթերինցիք , ընդհակառակը՝ նրա կողքով անցնելիս , ծափահարում էին , շվշվացնում և , ձեռք առնելով , ընդգծված բարձրագոչ ողջունում ու շնորհակալություն հայտնում .\n/.\n/interp/.\n ! "</p>	<p>Die Slytherins dagegen klatschten in die Hände , wenn er vorbeiging , sie piffen und johlten : " - - Danke , Potter , wir schulden dir noch was ! "</p>
<p>79231 Walking back from the library on his own one afternoon , he heard somebody whimpering from a classroom up ahead .</p>	<p>Կեսօրից հետո միայնակ վերադառնալով գրադարանից , նա ինչ-որ թնգթնգոց լսեց լսարաններից մեկի բաց դռնից :</p>	<p>Eines Nachmittags , auf dem Rückweg von der Bibliothek , hörte er in einem der Klassenzimmer vor ihm jemanden wimmern . - -</p>
<p>79432 All the same , he ' d have gambled twelve Sorcerer ' s Stones that Snape had just left the room , and from what Harry had just heard , Snape would be walking with a new spring in his step — Quirrell seemed to have given in at last .</p>	<p>Միևնույն ժամանակ , նա գրազ կգար տասներկու փրիստփայական քարերի վրա , որ Սառկեկը մի պահ առաջ դուրս եկավ լսարանի այդ կիսաբաց դռնով և դատելով իր լսածից , Հարրին համոզված էր , որ նա ուրախ-ուրախ թռչկոտելով էր հեռանում , որովհետև կարծես թե Քուիրելը վերջապես տեղի էր տվել նրա հնչմանը :</p>	<p>Dennoch : zwölf Steine der Weisen hätte er gewettet , daß es Snape war , der soeben das Zimmer verlassen hatte , und nach dem zu schließen , was Harry mitgehört hatte , gewiß mit federnden Schritten Quirrell schien nun doch nachgegeben zu haben .</p>
<p>81239 They walked in silence , their eyes on the ground .</p>	<p>Միառժամանակ նրանք լուռ քայլում էին , աչքերը գետնին գամած :</p>	<p>Sie gingen schweigend , die Augen auf die Erde gerichtet .</p>
<p>81329 They walked past a mossy tree stump .</p>	<p>Նրանք անցան մի մամուռակալած կոճղի կողքով :</p>	<p>Sie kamen an einem moosbewachsenen Baumstumpf vorbei .</p>
<p>81549 They walked more slowly , ears straining for the faintest sound .</p>	<p>Նրանք առաջ քայլեցին , սակայն ավելի դանդաղ և ականջներն աշալուրջ սրած անտառից եկող յուրաքանչյուր շշուկին :</p>	<p>Sie gingen jetzt langsamer , gespannt auf das leiseste Geräusch achtend .</p>
<p>81664 He walked forward and shook the centaur ' s hand .</p>	<p>Նա առաջ քայլեց ու շերմորեն սեղմեց կենտավրոսի ձեռքը :</p>	<p>Er trat vor und schüttelte die Hand des Zentauren .</p>
<p>82103 Bane walked over to stand next to Ronan .</p>	<p>Բեյնը մոտեցավ ու կանգնեց Ռոնանի կողքին :</p>	<p>Bane kam näher und stellte sich neben Ronan .</p>
<p>82337 They walked on through the dense , dark trees .</p>	<p>Նրանք առաջ շարժվեցին ծառերի մութ , խիտ բների միջով :</p>	<p>Sie gingen weiter durch dichten , dunklen Wald . -</p>
<p>82710 They walked for nearly half an hour , deeper and deeper into the forest , until the path became almost impossible to follow because the trees were so thick .</p>	<p>Մոտ կես ժամ նրանք լուռ քայլում էին , ավելի ու ավելի հեռանալով անտառի թավուտի մեջ , մինչև ծառերի բներն այնքան խտացան , որ արահետը գրեթե չէր երևում և առաջ գնալը գրեթե անհնար էր դարձել խիտ մացառուտների ու ծառերի ցածր՝ տերևաշատ հյուղերի պատճառով :</p>	<p>Sie gingen fast eine halbe Stunde lang tiefer und tiefer hinein , bis der Pfad sich fast verlor , so dicht standen die Bäume .</p>
<p>83613 Firenze slowed to a walk , warned Harry to keep his head bowed in case of low - hanging branches , but did not answer Harry ' s question .</p>	<p>Ֆիրենզին իր վազքը դանդաղեցրեց և Հարրիին զգուշացրեց , որ գլուխը ցած պահի՝ ծառերի ցածր հյուղերից շվսավելու համար , սակայն Հարրիի հարցերին չպատասխանեց :</p>	<p>Firenze ging nun im Schritt und ermahnte Harry , wegen der tiefen Äste den Kopf gesenkt zu halten , doch er antwortete nicht auf seine Fragen .</p>
<p>91182 He walked up to a black knight and put his hand out to touch the knight ' s horse .</p>	<p>Նա առաջ քայլեց՝ մոտենալով սև ձիուն և ձեռքը մեկնելով հպվեց սև քարին : -</p>	<p>Er ging vor zu einem schwarzen Springer , streckte die Hand aus und berührte ihn .</p>
<p>91403 The chessmen seemed to have been listening , because at these words a knight , a bishop , and a castle turned their backs on the white pieces and walked off the board , leaving three empty squares that Harry , Ron , and Hermione took .</p>	<p>Կարծես թե շախմատամարտիկները լսում էին , որովհետև այն պահին , երբ Ռոնը ավարտեց իր խոսքը մի ձի , մի փիղ և մի նավակ մեջքով շրջվեցին սպիտակների կողմը և դուրս քայլեցին շախմատային հարթակի վրայից՝ դատարկ թողնելով երեք շախմատավանդակ , որոնք էլ զբաղեցրին Ռոնը , Հարրին ու Հերմիոնան :</p>	<p>Die Schachfiguren hatten offenbar zugehört , denn in diesem Augenblick kehrten ein Springer , ein Läufer und ein Turm den weißen Figuren den Rücken und schritten vom Platz . Sie ließen drei leere Quadrate zurück , auf denen Harry , Ron und Hermine ihre Plätze einnahmen .</p>
<p>92573 Then she walked up and down the line of bottles , muttering to</p>	<p>Հետո նա ետ ու առաջ քայլեց շշերի շարքով՝ ինչ-որ բան մրմնջալով ու</p>	<p>Dann ging sie vor den Flaschen auf und ab , vor sich hin murmelnd</p>

Famous parallel texts

The Bible (1300+ languages)

The Universal Declaration of Human Rights (300+ languages)

The proceedings of the European Parliament (20+ languages)

Parallel corpora in comparative linguistics

Why are parallel texts interesting for linguists?

- translational equivalence
- available in many languages
- considered 'natural' language
- relatively easily attainable data

An example

The boy's name was Santiago. Dusk was falling as the boy arrived with his herd at an abandoned church. The roof had fallen in long ago, and an enormous sycamore had grown on the spot where the sacristy had once stood.

He decided to spend the night there. He saw to it that all the sheep entered through the ruined gate, and then laid some planks across it to prevent the flock from wandering away during the night. There were no wolves in the region, but once an animal had strayed during the night, and the boy had had to spend the entire next day searching for it.

He swept the floor with his jacket and lay down, using the book he had just finished reading as a pillow. He told himself that he would have to start reading thicker books: they lasted longer, and made more comfortable pillows.

It was still dark when he awoke, and, looking up, he could see the stars through the half-destroyed roof.

I wanted to sleep a little longer, he thought. He had had the same dream that night as a week ago, and once again he had awakened before it ended.

He arose and, taking up his crook, began to awaken the sheep that still slept. He had noticed that, as soon

Erster Teil

I

Der Jüngling hieß Santiago. Es fing bereits an zu dämmern, als er mit seiner Schafherde zu einer alten, verlassenem Kirche kam. Das Dach war schon vor geraumer Zeit eingestürzt, und ein riesiger Feigenbaum wuchs an jener Stelle, an der sich einst die Sakristei befand.

Er entschloß sich, die Nacht hier zu verbringen. Er trieb alle Schafe durch die zerfallene Tür und legte einige Bretter so davor, daß ihm die Tiere während der Nacht nicht entweichen konnten. Zwar gab es keine Wölfe in jener Gegend, aber einmal war ihm eines der Tiere während der Nacht entkommen, und er hatte den ganzen folgenden Tag mit der Suche nach dem verirrtten Schaf verbracht.

Dann breitete er seinen Mantel auf dem Fußboden aus, legte sich nieder und nahm das Buch, in dem er gerade gelesen hatte, als Kopfkissen. Vor dem Einschlafen überlegte er sich, daß er in Zukunft dickere Bücher lesen sollte, weil man länger etwas davon hatte und weil sie eine bequemere Kopfstütze abgaben.

Es war noch finster, als er erwachte. Als er nach oben schaute, sah er die Sterne zwischen den Dachbalken durchscheinen.

›Eigentlich wollte ich noch weiterschlafen‹, dachte er bei sich. Wieder hatte er den gleichen Traum gehabt wie

An example

The boy's name was Santiago. Dusk was falling as the boy arrived with his herd at an abandoned church. The roof had fallen in long ago, and an enormous sycamore had grown on the spot where the sacristy had once stood.

He decided to spend the night there. He saw to it that all the sheep entered through the ruined gate, and then laid some planks across it to prevent the flock from wandering away during the night. There were no wolves in the region, but once an animal had strayed during the night, and the boy had had to spend the entire next day searching for it.

He swept the floor with his jacket and lay down, using the book he had just finished reading as a pillow. He told himself that he would have to start reading thicker books: they lasted longer, and made more comfortable pillows.

It was still dark when he awoke, and, looking up, he could see the stars through the half-destroyed roof.

I wanted to sleep a little longer, he thought. He had had the same dream that night as a week ago, and once again he had awakened before it ended.

He arose and, taking up his crook, began to awaken the sheep that still slept. He had noticed that, as soon

Erster Teil

I

Der Jüngling hieß Santiago. Es fing bereits an zu dämmern, als er mit seiner Schafherde zu einer alten, verlassenem Kirche kam. Das Dach war schon vor geraumer Zeit eingestürzt, und ein riesiger Feigenbaum wuchs an jener Stelle, an der sich einst die Sakristei befand.

Er entschloß sich, die Nacht hier zu verbringen. Er trieb alle Schafe durch die zerfallene Tür und legte einige Bretter so davor, daß ihm die Tiere während der Nacht nicht entweichen konnten. Zwar gab es keine Wölfe in jener Gegend, aber einmal war ihm eines der Tiere während der Nacht entkommen, und er hatte den ganzen folgenden Tag mit der Suche nach dem verirrt Schaf verbracht.

Dann breitete er seinen Mantel auf dem Fußboden aus, legte sich nieder und nahm das Buch, in dem er gerade gelesen hatte, als Kopfkissen. Vor dem Einschlafen überlegte er sich, daß er in Zukunft dickere Bücher lesen sollte, weil man länger etwas davon hatte und weil sie eine bequemere Kopfstütze abgaben.

Es war noch finster, als er erwachte. Als er nach oben schaute, sah er die Sterne zwischen den Dachbalken durchscheinen.

›Eigentlich wollte ich noch weiterschlafen‹, dachte er bei sich. Wieder hatte er den gleichen Traum gehabt wie

An example

The boy's name was Santiago. Dusk was falling as the boy arrived with his herd at an abandoned church. The roof had fallen in long ago, and an enormous sycamore had grown on the spot where the sacristy had once stood.

He decided to spend the night there. He saw to it that all the sheep entered through the ruined gate, and then laid some planks across it to prevent the flock from wandering away during the night. There were no wolves in the region, but once an animal had strayed during the night, and the boy had had to spend the entire next day searching for it.

He swept the floor with his jacket and lay down, using the book he had just finished reading as a pillow. He told himself that he would have to start reading thicker books: they lasted longer, and made more comfortable pillows.

It was still dark when he awoke, and, looking up, he could see the stars through the half-destroyed roof.

I wanted to sleep a little longer, he thought. He had had the same dream that night as a week ago, and once again he had awakened before it ended.

He arose and, taking up his crook, began to awaken the sheep that still slept. He had noticed that, as soon

Erster Teil

I

Der Jüngling hieß Santiago. Es fing bereits an zu dämmern, als er mit seiner Schafherde zu einer alten, verlassenem Kirche kam. Das Dach war schon vor geraumer Zeit eingestürzt, und ein riesiger Feigenbaum wuchs an jener Stelle, an der sich einst die Sakristei befand.

Er entschloß sich, die Nacht hier zu verbringen. Er trieb alle Schafe durch die zerfallene Tür und legte einige Bretter so davor, daß ihm die Tiere während der Nacht nicht entweichen konnten. Zwar gab es keine Wölfe in jener Gegend, aber einmal war ihm eines der Tiere während der Nacht entkommen, und er hatte den ganzen folgenden Tag mit der Suche nach dem verirrtten Schaf verbracht.

Dann breitete er seinen Mantel auf dem Fußboden aus, legte sich nieder und nahm das Buch, in dem er gerade gelesen hatte, als Kopfkissen. Vor dem Einschlafen überlegte er sich, daß er in Zukunft dickere Bücher lesen sollte, weil man länger etwas davon hatte und weil sie eine bequemere Kopfstütze abgaben.

Es war noch finster, als er erwachte. Als er nach oben schaute, sah er die Sterne zwischen den Dachbalken durchscheinen.

›Eigentlich wollte ich noch weiterschlafen‹, dachte er bei sich. Wieder hatte er den gleichen Traum gehabt wie

Parallel corpora in comparative linguistics

Stolz (2005, 2006):

‘Le Petit Prince’ in 64 languages

comitatives and instrumentals

“Then he mopped his forehead **with** a
handkerchief decorated **with** red
squares.”

Parallel corpora in comparative linguistics

Van der Auwera et al. (2005):

‘Harry Potter and the chamber of
secrets’ in 10 Slavic languages

expression of uncertainty: the use of
verbs like ‘may’, ‘might’, and ‘could’
versus that of adverbs like ‘maybe’ and
‘perhaps’.

Parallel corpora in comparative linguistics

Wälchli (2009):

The 'Gospel according to Mark' in 100+
languages

Lexicalisation in motion events

The use of different types of motion
verbs seems not to be determined by
genetic relationships between
languages, but by areal factors

Parallel corpora in comparative linguistics

My own corpus:

Alice's adventures in Wonderland /
Through the Looking-Glass and what
Alice found there (Lewis Carroll) /
O Alquimista (Paulo Coelho) in 20+
languages

Syntactic and semantic change in
motion event encoding in the Indo-
European language family

Advantages

- usage-based rather than typifying
- once properly build, can be used for the investigation of many different topics
- comparability of original and translation is helpful for data analysis

Disadvantages

- translations into non-European languages are less common and harder to find
- the translation might be distorted because of the source text
- written language instead of spoken language

Non-comparative uses of parallel corpora

deciphering ancient texts

machine translation technology

Conclusion

Parallel corpora are a great resource for comparative linguists

More online accessible parallel corpora would provide a great resource

Thank you!

annemarie.verkerk@mpi.nl