

Dutch Parallel Corpus

Lieve Macken

lieve.macken@hogent.be

LT³, Language and Translation Technology Team
Faculty of Applied Language Studies
University College Ghent

November 29th 2011





Dutch Parallel Corpus

Characteristics

- Parallel Corpus of 10 million words
- 2 language pairs, 4 language directions
 - Dutch–French / Dutch–English
 - Bidirectional
 - Also comparable corpus (translated vs. non-translated language)
- Balanced
 - 5 text types: administrative texts, instructive texts, literature, journalistic texts, texts for external communication
 - Text types further divided into subtypes
 - \pm 500,000 words per translation direction/text type
 - Metadata files
 - info on translation direction and text types
 - other relevant info such as intended audience, text provider, ...



Dutch Parallel Corpus

Characteristics (ctd.)

- Balanced (ctd.)
 - Condition on translation direction relaxed for instructive texts
 - difficult to find information on translation direction
 - Literary texts balanced according to language pairs
 - difficult to obtain copyright clearance
- Copyright-cleared
 - Difficult and time-consuming task
 - Commercial publishers (journalistic texts and literature) → longer and more difficult negotiation cycles
 - Different type of IPR-agreements
 - commercial vs. non-commercial use
 - recognizable vs. partially recognizable



Characteristics (ctd.)

- Automatically aligned at sentence level
 - Combined three alignment tools
 - reduce manual verification effort
 - No word alignment
- Linguistically enriched
 - Preprocessing
 - Character conversion (UTF-8)
 - Text split in sentences and tokens
 - Part-of-speech tagging
 - Info on tag sets: see user doc
 - Lemmatization



Dutch Parallel Corpus

Characteristics (ctd.)

Obtained accuracy scores on a manually validated DPC sample

	Sample size (tokens)	Lemmata	PoS (full tag)	PoS (main category)
Dutch	211,000	96.5%	94.8%	97.4%
English	300,000	98.1%	96.2%	N/A
French	330,000	98.1%	94.6%	97.4%



Distribution

- Distributed and maintained by HLT-agency (TST-centrale)
 - <http://www.inl.nl/tst-centrale/nl/producten/corpora/dpc-corpus-dutch-parallel-corpus/6-65>



4 files per text

- 2 monolingual files
 - Indications of paragraphs and sentences
 - Linguistic annotations
- File containing metadata
 - Author, date, publication, ...
 - Source and target language
 - Text type and subtype
 - Domain, keywords
 - Type of IPR-agreement
- File with sentence alignments



Monolingual XML file

```
<p n="p1">
<seg type="sent" n="seg.p1.s1">
<seg type="original">Niet alleen voor wereldverbeteraars</seg>
<s n="p1.s1">
<w ana="BW()" lemma="niet">Niet</w>
<w ana="BW()" lemma="alleen">alleen</w>
<w ana="VZ(init)" lemma="voor">voor</w>
<w ana="N(soort,mv,basis)" lemma="wereldverbeteraar">wereldverbeteraars</w></s>
</seg>
<seg type="sent" n="seg.p1.s2">
<seg type="original">De 7 misverstanden over ethisch beleggen</seg>
<s n="p1.s2">
<w ana="LID(bep,stan,rest)" lemma="de">De</w>
<w ana="TW(hoofd,prenom,stan)" lemma="7">7</w>
<w ana="N(soort,mv,basis)" lemma="misverstand">misverstanden</w>
<w ana="VZ(init)" lemma="over">over</w>
<w ana="ADJ(prenom,basis,zonder)" lemma="ethisch">ethisch</w>
<w ana="WW(inf,vrij,zonder)" lemma="beleggen">beleggen</w></s>
</seg>
```



Metadata XML file

```
<Outcome>written to be read</Outcome>  
<TextType>Journalistic texts</TextType>  
<TextSubType>Comment articles (background articles)</TextSubType>  
<Domain>Finance</Domain>  
<Keywords1>Banking</Keywords1>  
<IPR>light version</IPR>  
<TypeOfInstitution>profit</TypeOfInstitution>  
<IntendedAudience>broad external audience</IntendedAudience>  
</metaText>  
<metaTrans>  
<Original lang= "NL" />  
<Translated>FR</Translated>  
<Intermediate>not applicable</Intermediate>  
<Mode>human</Mode>  
</metaTrans>
```



Alignment XML file

```
<linkGrp type="alignment">
<link type="B: 1-1" targets="p1.s1; p1.s1"/>
<link type="B: 1-1" targets="p1.s2; p1.s2"/>
<link type="B: 1-1" targets="p1.s3; p1.s3"/>
<link type="B: 1-1" targets="p1.s4; p1.s4"/>
<link type="A: 2-1" targets="p1.s5 p1.s6; p1.s5"/>
<link type="B: 1-1" targets="p1.s7; p1.s6"/>
<link type="A: 1-1" targets="p1.s8; p1.s7"/>
<link type="A: 1-1" targets="p1.s9; p1.s8"/>
<link type="A: 1-1" targets="p1.s10; p1.s9"/>
<link type="B: 1-1" targets="p1.s11; p1.s10"/>
<link type="A: 1-1" targets="p1.s12; p1.s11"/>
<link type="A: 1-1" targets="p1.s13; p1.s12"/>
<link type="A: 1-1" targets="p1.s14; p1.s13"/>
<link type="A: 1-1" targets="p1.s15; p1.s14"/>
<link type="A: 1-1" targets="p1.s16; p1.s15"/>
<link type="A: 1-1" targets="p1.s17; p1.s16"/>
```

Webinterface




Demo version

- <http://dpc.inl.nl/indexd.php>



DEMOVERSIE MET BEPERKTE MOGELIJKHEDEN

DPC SEARCH

Klik hier voor de handleiding 

Eéntalig zoeken

taal:

Tweetalig zoeken

taal1: en taal2:

doorzoek alle teksten - tekstselectie

Webinterface ctd.



Three steps

- ① Monolingual vs. bilingual search
- ② Full corpus vs. subcorpus
- ③ Search query



Putting together a subcorpus

- Multifunctional resource for diverse group of users
 - Researchers in translation studies vs. contrastive linguistics
 - Language teachers and language learners
 - Human translators
- Create a subcorpus according to your needs!
 - E.g. filter on translation direction
 - E.g. filter on text type
 - But: the more restrictions . . . the less data

Webinterface ctd.



DPC SEARCH

Klik hier voor de handleiding 

Eéntalig zoeken

taal:

Tweetalig zoeken

taal1: en taal2:

doorzoek alle teksten - tekstselectie

subcorpus

teksttype <input type="text" value="Journalistic texts"/>	domein <input type="text" value="kies hier"/>	copyright <input type="text" value="kies hier"/>	brontaal <input type="text" value="EN"/>	beschikbare talen <input type="text" value="NL"/>
--	--	---	---	--

Het subcorpus van taal 1 bestaat uit **305677** woorden.
Het subcorpus van taal 2 bestaat uit **295039** woorden.



Search query

- Define search words
 - max. 2 per language
- Define search type
 - Word form or lemma
 - Part-of-speech
 - Position



Example 2 consecutive words

woord1

du op woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort

Maak uw keuze ▾

positie

de positie is niet belangrijk ▾

EN NIET

woord2

coup op woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort

Maak uw keuze ▾

positie

de positie is niet belangrijk ▾

maximum aantal ▾ woorden tussen de zoekwoorden:



Example 2 consecutive words (results)

i C	Du coup , le match était joué.	i C	Daarmee was de wedstrijd gespeeld.
i C	C'étaient de bonnes campagnes pour l'époque, pondérées et, du coup , cela traduisait tout le sérieux du personnage, qui a de l'envergure.	i C	Het waren voor hun tijd goede en evenwichtige campagnes die meteen ook illustratief waren voor de sérieux van de persoon die charisma heeft.
i C	Du coup , chacun se précipita pour profiter de cette chance imprévue et inexplicable.	i C	René van Lotharingen keerde naar zijn land terug. In Gelre brak een opstand uit.
i C	« Souvent, nous nous projetons dans le futur ou nous ressasons le passé. Et du coup , le moment que nous vivons nous échappe.	i C	"Vaak zijn we bezig met toekomstige problemen of analyseren we het verleden, waardoor het enige moment waarin we nu zijn, ons ontglipt.
i C	Du coup , j'ai cherché à présenter un grand nombre de pages sur le site Web.	i C	Daarom geef ik zoveel mogelijk informatie via de website.
i C	Du coup , le téléphone a pris position dans la cuisine.	i C	Daarom werd de telefoon al snel in de keuken geplaatst.
i C	Du coup , on faisait des efforts persévérants pour réparer les dégâts, reconstruire l'économie, extraire des conditions du Traité tous les avantages possibles.	i C	Onmiddellijk werden er aanhoudend inspanningen ondernomen om de schade te herstellen, de economie opnieuw op te bouwen en zo veel mogelijk voordeel te halen uit de beperkte mogelijkheden van het verdrag.
i C	Du coup , les méridionaux perdirent toutes les illusions qu'ils étaient en droit d'avoir sur leur situation dans le nouvel Etat.	i C	Zo verdween iedere illusie die het zuiden had kunnen koesteren over zijn positie in de nieuwe staat.
i C	Dans le sud, la répression de la contre-réforme, la censure et la crainte étouffèrent l'activité intellectuelle, tandis que la langue néerlandaise était redoutée en tant qu'arme du protestantisme; du coup on ne fit rien pour maintenir ni pour resserrer les contacts littéraires avec le nord.	i C	In het zuidelijke deel van de Nederlanden verstikten de onderdrukking van de Contrareformatie, de censuur en de angst de intellectuele activiteiten, terwijl de Nederlandse taal zelf gevreesd werd als het wapen van het protestantisme en daarom werd er ook niets ondernomen om de literaire banden met het noorden in stand te houden of te versterken.
i C	En supprimant la vodka, on a adouci la boisson, qui du coup plaît plus aux femmes.	i C	Wodka is 'harde' alcohol en die wordt meer geassocieerd met mannen.



Webinterface ctd.

Metadata

1	Filename	dpc-rou-000389
2	Language	FR (BE)
3	Author/translator	Bruno Govers
4	Text unit title	Zvonko Mijevic et Peter Schmeichel
5	magazine/journal title	Sport Magazine
6	publisher	Roularta
19	Text type	Journalistic texts
20	Text subtype	Comment articles (background articles)
21	Domain	Communication
22	Keywords	Sports
23	Copyright/IPR-agreement	light version
24	Type of institution	profit
25	Intended audience	broad external audience
26	Original text & language	FR
27	Translated text & language	NL
28	Intermediate text & language	not applicable
29	Translation mode	unknown
30	Number of words	635
31	Number of sentences	28

Context

Toon een context van zinnen.

Jusqu'à l'heure, mes coéquipiers et moi avons fait davantage que jeu égal avec les "Red Devils". A la 67e minute, toutefois, ceux-ci bénéficièrent d'un penalty généreux.

En première instance, j'étais parvenu à dévier la tentative du capitaine des "Mancunians", Steve Bruce mais le rebond fut favorable à son partenaire, Brian Mc Clair.

Du coup, le match était joué.

J'avais 20 ans et venais d'accéder au statut de titulaire à l'Etoile Rouge Belgrade.

Le Danois, lui, avait déjà un peu plus de bouteille, car avant d'effectuer ses débuts à Manchester, il faisait déjà figure depuis longtemps d'incontournable à Brøndby ainsi qu'au sein de l'équipe nationale de son pays.

Compte tenu de son impressionnant physique, je m'étais fait la réflexion qu'il était taillé sur mesure pour le football anglais.



Export to Excel

- Source and target sentences and accompanying metadata

taal1	taal2	Language - taal1	Text type - taal1	Text subtype - taal1
Du coup, le match était joué. C'étaient de bonnes campagnes pour l'époque, pondérées et, du coup, cela traduisait tout le sérieux du personnage, qui a de l'envergure.	Daarmee was de wedstrijd gespeeld. Het waren voor hun tijd goede en evenwichtige campagnes die meteen ook illustratief waren voor de sérieux van de persoon die charisma heeft.	FR (BE) FR (BE)	Journalistic texts Non-fictional literature	Comment articles (background articles) (Auto)biographies
Du coup, chacun se précipita pour profiter de cette chance imprévue et inexplicable. « Souvent, nous nous projetons dans le futur ou nous rressas-sons le passé. Et du coup, le moment que nous vivons nous échappe.	René van Lotharingen keerde naar zijn land terug. In Gelre brak een opstand uit. "Vaak zijn we bezig met toekomstige problemen of analyseren we het verleden, waardoor het enige moment waarin we nu zijn, ons ontglipt.	FR (FR) FR (BE)	Non-fictional literature Journalistic texts	Expository works of a general nature Comment articles (background articles)
Du coup, j'ai cherché à présenter un grand nombre de pages sur le site Web.	Daarom geef ik zoveel mogelijk informatie via de website.	FR (BE)	Journalistic texts	Comment articles (background articles)



Puisque – aangezien

<p>woord1</p> <p><input type="text" value="puisque"/> op <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p><input type="text" value="Maak uw keuze"/></p> <p>positie</p> <p><input type="text" value="de positie is niet belangrijk"/></p>	<p>alles wissen</p>	<p>woord1</p> <p><input type="text" value="aangezien"/> op <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p><input type="text" value="Maak uw keuze"/></p> <p>positie</p> <p><input type="text" value="de positie is niet belangrijk"/></p>
<p><input checked="" type="radio"/> EN <input type="radio"/> NIET</p>		<p><input checked="" type="radio"/> EN <input type="radio"/> NIET</p>
<p>woord2</p> <p><input type="text"/> op <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p><input type="text" value="Maak uw keuze"/></p> <p>positie</p> <p><input type="text" value="de positie is niet belangrijk"/></p>		<p>woord2</p> <p><input type="text"/> by <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p><input type="text" value="Maak uw keuze"/></p> <p>positie</p> <p><input type="text" value="de positie is niet belangrijk"/></p>
<p><input type="text" value="maximum aantal"/> woorden tussen de zoekwoorden: <input type="text"/></p>		<p><input type="text" value="maximum aantal"/> woorden tussen de zoekwoorden: <input type="text"/></p>



Results Puisque – aangezien

<p>i c Ce montant avait été revu à la baisse lors de la conclusion de l'accord puisque le budget 2007 décidé auparavant prévoit 1.523 milliers €.</p>	<p>i c Toen het akkoord werd gesloten, werd dit bedrag verminderd, aangezien er in het budget 2007 dat aanvankelijk werd beslist, een bedrag van 1.523 miljoen € was uitgetrokken.</p>
<p>i c Cette modification permet une simplification administrative puisque le patient ne devra plus présenter de preuves de paiement pour bénéficier d'un remboursement.</p>	<p>i c Door die wijziging wordt een administratieve vereenvoudiging mogelijk gemaakt, aangezien de patiënt geen betalingsbewijzen meer zal moeten voorleggen om een terugbetaling te genieten.</p>
<p>i c On ne peut en effet ici imaginer une quelconque prévention des maladies génétiques puisque l'empreinte génétique à des fins d'identification est en effet pratiquée sur de l'ADN non codant à partir duquel aucune maladie génétique ne saurait être détectée.</p>	<p>i c Het is immers onvoorstelbaar dat hierdoor erfelijke ziektes zouden kunnen worden voorkomen aangezien het onderzoek naar de genetische vingerafdruk voor identificatiedoelinden wordt uitgevoerd op niet-coderend DNA, waarin geen enkele erfelijke ziekte kan worden opgespoord.</p>
<p>i c La C.C.B. constate que la première mesure proposée n'a pas d'impact sur le budget puisque les patients visés sont déjà remboursés dans le cadre actuel... .</p>	<p>i c De C.B.C. stelt vast dat de eerste maatregel geen weerslag heeft op het budget, aangezien de beoogde patiënten al in het huidige kader worden terugbetaald.. .</p>
<p>i c " Monsieur Arickx a proposé une feuille de calcul pour qu'il n'y ait pas d'impact sur le patient puisque, dans cette proposition, le ticket modérateur reste le même par produit.</p>	<p>i c " De heer Arickx heeft een berekeningsblad voorgesteld, opdat er geen weerslag voor de patiënt zou zijn, aangezien in dat voorstel het remgeld per product hetzelfde blijft.</p>



Dutch compounds

woord1
?*maatregel* op woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort
Maak uw keuze ▾

positie
de positie is niet belangrijk ▾

EN NIET

woord2
 by woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort
Maak uw keuze ▾

positie
de positie is niet belangrijk ▾

maximum aantal ▾ woorden tussen de zoekwoorden:

alles wissen

woord1
 op woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort
adjective ▾ [sluiten]
Maak uw keuze ▾

positie
de positie is niet belangrijk ▾

EN NIET

woord2
measure* op woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort
Maak uw keuze ▾

positie
de positie is niet belangrijk ▾

maximum aantal ▾ woorden tussen de zoekwoorden:



Results Dutch compounds

<p>i c Daar hoort noodzakelijkerwijs een goed functionerend controlesysteem met strenge strafmaatregelen bij.</p>	<p>i c This must also include a properly functioning control system with strict punitive measures.</p>
<p>i c Vandaag ziet het er steeds meer naar uit dat de economische groeivertraging geleidelijk aan het uitbodemen is en dat de economie opnieuw in een opwaartse fase zal terechtkomen, zeker na de fiscale en monetaire stimuleringsmaatregelen in de nasleep van 11 september.</p>	<p>i c It is looking increasingly as though the economic trough is bottoming out and that the economy will start to pick up again, especially following the fiscal and monetary measures adopted in the aftermath of the 11th of September.</p>
<p>i c De Raad heeft er zich voor wat betreft de structuurmaatregelen, rubriek 2 van de financiële vooruitzichten, aan gehouden deze gehele rubriek in de begroting op te nemen als kredieten voor vastlegging overeenkomstig de conclusies van de Europese Raad van Berlijn van maart 1999.</p>	<p>i c With regard to the structural measures, heading 2 of the financial perspectives, the Council insisted on including this entire heading in the budget in the form of commitment appropriations in accordance with the conclusions of the European Council of Berlin of March 1999.</p>
<p>i c Hoe slecht de Raad op cruciale momenten functioneert, is wel gebleken bij de Top van Gent, waar drie minister-presidenten vooraf afspraken maakten buiten de Raad om, waar de voorzitter van de Raad essentiële uitslatingen deed over de evaluatie van de aanvallen op Afghanistan die later niet door de Raad werden gedekt, waar diverse minister-presidenten verschillende uitslatingen deden na de Europese top, waar besluiten werden genomen over een betalingskwesitie van banken met betrekking tot de euro, maar geen besluiten vielen over cruciale door Commissie en Parlement voorgestelde antiterrorismemaatregelen en waar tenslotte ook nog een gênante ruzie ontstond tussen de voorzitter van de Raad en de voorzitter van de Commissie.</p>	<p>i c How badly the Council functions at crucial moments was evident at the Ghent Summit where three Prime Ministers had made prior arrangements outside of the Council, where the President of the Council made essential statements on the evaluation of the attacks on Afghanistan which were subsequently not backed by the Council, where various Prime Ministers made various statements following the European Summit, where decisions were taken on a payment issue of banks in relation to the euro, but where no decisions were made on crucial anti-terrorist measures proposed by the Commission and Parliament and where, to cap it all, an embarrassing dispute arose between the President of the Council and the President of the Commission.</p>
<p>i c Commissaris Vitorino heeft ons een overzicht gegeven met een schets van dergelijke antiterrorismemaatregelen.</p>	<p>i c Commissioner Vitorino has provided us with an anti-terrorism road map outlining such measures.</p>



Passé composé (French) – Simple past (Dutch)

<p>woord1</p> <p>avoir <input type="radio"/> op <input checked="" type="radio"/> woordvorm <input checked="" type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsort</p> <p>verbe <input type="text"/> <input checked="" type="radio"/> tous les verbes <input type="radio"/> verbe principal <input checked="" type="radio"/> verbe [sluiten]</p> <p>auxiliaire</p> <p>indicatif <input type="text"/></p> <p><input checked="" type="radio"/> sing. + plur. <input type="radio"/> singulier <input type="radio"/> pluriel</p> <p><input checked="" type="radio"/> 1e + 2e + 3e pers. <input type="radio"/> 1e pers. <input type="radio"/> 2e pers. <input type="radio"/> 3e pers.</p> <p><input type="radio"/> tous les temps <input checked="" type="radio"/> présent <input type="radio"/> imparfait <input type="radio"/> futur simple <input type="radio"/> passé simple</p> <p>positie</p> <p>de positie is niet belangrijk <input type="text"/></p> <p><input checked="" type="radio"/> EN <input type="radio"/> NIET</p>	<p>alles wissen</p>	<p>woord1</p> <p><input type="text"/> op <input type="radio"/> woordvorm <input checked="" type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsort</p> <p>werkwoord <input type="text"/> [sluiten]</p> <p>persoonsvorm <input type="text"/></p> <p><input checked="" type="radio"/> alle vormen <input type="radio"/> enkelvoud <input type="radio"/> meervoud <input type="radio"/> met -t</p> <p><input type="radio"/> alle tijden <input type="radio"/> tegenwoordige tijd <input checked="" type="radio"/> verleden tijd <input type="radio"/> conjunctief</p> <p>positie</p> <p>de positie is niet belangrijk <input type="text"/></p> <p><input checked="" type="radio"/> EN <input type="radio"/> NIET</p>
<p>woord2</p> <p><input type="text"/> op <input type="radio"/> woordvorm <input checked="" type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsort</p> <p>verbe <input type="text"/> <input checked="" type="radio"/> tous les verbes <input type="radio"/> verbe principal <input type="radio"/> verbe [sluiten]</p> <p>auxiliaire</p> <p>participe passé <input type="text"/></p> <p><input checked="" type="radio"/> masc. + fém. <input type="radio"/> masculin <input type="radio"/> féminin</p> <p><input checked="" type="radio"/> sing. + plur. <input type="radio"/> singulier <input type="radio"/> pluriel</p> <p>positie</p> <p>de positie is niet belangrijk <input type="text"/></p> <p>maximum aantal <input type="text"/> woorden tussen de zoekwoorden: 0</p>		<p>woord2</p> <p><input type="text"/> by <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsort</p>



Results passé composé (French) – Simple past (Dutch)

<p>1 C « Nous avons un bon produit et une gamme étendue, et nous leur avons donné une marque forte.</p>	<p>1 C "We hadden een goed product en een breed gamma en hebben er een sterk merk op geplakt.</p>
<p>1 C Il à expliqué que dans son pays, l'accès aux écoles publiques est prévu pour tous y compris les illégaux.</p>	<p>1 C Hij legde uit dat in zijn land iedereen tot het openbaar onderwijs wordt toegelaten, ook illegalen.</p>
<p>1 C Nous leur avons annoncé ensemble que nous allions dévaluer.</p>	<p>1 C We deelden hun mee dat we gingen devalueren.</p>
<p>1 C Nous avons voulu et avons réussi à développer en la matière une véritable compétence.</p>	<p>1 C Wij wensten op dat vlak een grote knowhow te ontwikkelen en zijn daar ook in geslaagd.</p>
<p>1 C Nous avons travaillé trois ans et demi ensemble, et avons quitté la boîte au même moment pour des raisons identiques: nous ne nous reconnaissons plus dans le métier, dans la façon trop égocentrique dont les grands bureaux envisageaient la publicité.</p>	<p>1 C Wij werkten drieënhalf jaar samen en verlieten samen het bedrijf om dezelfde reden: we voelden ons niet meer goed in onze job, we hadden het moeilijk met de egocentrische manier waarop de grote bureaus reclame benaderen.</p>
<p>1 C Lorsqu'un jeune âgé de 16 à 18 ans n'est pas ouvert à une approche pédagogique et lorsqu'il à soit commis des faits graves, soit déjà fait l'objet antérieurement d'une mesure protectionnelle, le juge de la jeunesse peut décider de demander des expertises en vue du dessaisissement.</p>	<p>1 C In geval een jongere, tussen 16 en 18 jaar oud, niet openstaat voor een pedagogische aanpak en hij ofwel ernstige feiten heeft gepleegd ofwel reeds voorheen een jeugdbeschermingsmaatregel kreeg opgelegd, kan de jeugdrechter beslissen deskundigen aan te wijzen met het oog op ulthandengeving.</p>
<p>1 C Je n'en ai tiré aucune vanité.</p>	<p>1 C Toch was dit voor mij geen reden om ijdel te worden.</p>
<p>1 C L'exemple le plus emblématique est évidemment le holding IRIS Group, qui a été mis sur pied en 1992 lorsque Pierre De Muelenaere et moi-même avons monté une opération de management buy-out (MBO) pour racheter IRIS, spécialisée dans la reconnaissance optique de caractères, au groupe Electrorail-Prodata.</p>	<p>1 C Het meest symbolische voorbeeld is uiteraard de holdingmaatschappij IRIS Group, opgericht in 1992 toen Pierre De Muelenaere en ikzelf een management-buy-out (MBO) op touw hadden gezet om IRIS, gespecialiseerd in optische karakterherkenning, van de groep Electrorail-Prodata over te nemen.</p>



Final verbal group in Dutch

woord1

op woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort

[sluiten]

alle vormen enkelvoud meervoud met -t

alle tijden tegenwoordige tijd verleden tijd conjunctief

positie

EN NIET

woord2

worden by woordvorm lemma (* is jokerteken, ? is één karakter jokerteken)

woordsoort

[sluiten]

positie



Results final verbal group in Dutch

- i C Brabantia biedt een vuilnisbak aan waarmee het afval samengedrukt **kan worden**.
- i C De banen in de dienstverleningscentra hadden helemaal niet hetzelfde prestige als het werk in kleine busi-nessunits, waar aangepaste oplossingen ontwikkeld **konden worden**.
- i C De Dienst houdt vol dat de vragen die hierover werden gesteld tijdens de eerste maanden van de omstreden periode talrijker waren dan wat op basis van het reële statistische aandeel van de stottergevallen in de volledige logopedische verzorging verwacht **kon worden**.
- i C Dit wil zeggen dat de jongere onmiddellijk, maar wel voorlopig, een aantal maatregelen opgelegd kan krijgen, ook al heeft de rechter nog niet geoordeeld over de schuld van de jongere en over de maatregel die in dat kader tegen hem genomen **moet worden**.
- i C Om het sociale leven te vrijwaren zouden gehoor- en tandprothesen beter gefinancierd **moeten worden**.
- i C Hij voegt eraan toe dat indien de betrokken personen in de toekomst meer arbeidsmogelijkheden krijgen, deze doelgroep op termijn vanzelf kleiner **zal worden**.
- i C De OISZ worden aangemoedigd hun meetsysteem te blijven ontwikkelen, waarbij een of meerdere indicatoren in de bestuursovereenkomst als operationele doelstelling opgenomen **kunnen worden**.
- i C Samen met de vier andere sorteercentra vormt het een samenhangende infrastructuur waarmee voortdurend een zeer hoog prestatieniveau gehaald **kan worden**.



Acknowledgements

- Funded by the STEVIN programme
- DPC Team
 - K.U. Leuven campus Kortrijk
 - Faculty of Applied Language Studies – University College Ghent
 - Contributors
 - Piet Desmet, Willy Vandeweghe, Hans Paulussen, Lieve Macken, Maribel Montero Perez, Orphée De Clercq, Lidia Rura, Julia Trushkina and Antoine Besnehard
- <http://www.kuleuven-kortrijk.be/dpc>
 - User Manual
 - Publications