

CLARIN ERIC

Toegevoegde waarde voor
onderzoek en ontwikkeling
in Vlaanderen

CLARIN ERIC:

Toegevoegde waarde voor onderzoek en
ontwikkeling in Vlaanderen

Nederlandse Taalunie

Instituut voor de Nederlandse Taal

Consortium CLARIN DLU/Flanders

CLARIN ERIC

INHOUDSOPGAVE

Inleiding.....	3
CLARIN: een meerwaarde voor Vlaamse onderzoekers en Vlaanderen.....	6
A. Voordelen voor onderzoekers	6
i. Raadplegen van data	6
ii. Archiveren/citeren van data.....	6
iii. Bewerken van data	7
iv. Eenmalige aanmelding.....	7
v. Toegang tot kennis.....	7
vi. Centrale diensten.....	8
vii. Deelname in Europese projecten	8
B. Voordelen voor Vlaanderen	8
CLARIN ERIC Vlaanderen 2012-2017.....	10
Onderzoek Taal- en Spraaktechnologie en Digital Humanities in Vlaanderen 2012-2017.....	14
A. KU Leuven – Overzicht projecten	14
B. Universiteit Antwerpen – Overzicht projecten	58
C. Universiteit Gent – Overzicht projecten	72
D. Vrije Universiteit Brussel – Overzicht projecten	81
Netwerken, valorisatie en maatschappelijke benutting.....	85
Toekomstige ontwikkelingen: CLARIAH-VL.....	87
Appendix: overzicht publicaties.....	89
A. KU Leuven – Overzicht publicaties	89
B. Universiteit Antwerpen - Overzicht publicaties	123
C. Universiteit Gent - Overzicht publicaties	135
D. Vrije Universiteit Brussel - Overzicht publicaties	144

INLEIDING

In dit rapport wordt een overzicht gegeven van het onderzoek op het gebied van taal- en spraaktechnologie en gerelateerde disciplines in Vlaanderen in de afgelopen vijf jaar. Het doel hiervan is een beeld te schetsen van de wetenschappelijke context waarin de activiteiten van CLARIN ERIC in Vlaanderen hebben plaatsgevonden en van het onderzoek dat Vlaamse deelname aan CLARIN ERIC motiveert en legitimeert.

Anders dan in de meeste landen die lid zijn van het CLARIN-ERIC-netwerk, is er in Vlaanderen geen sprake van een nationaal programma voor CLARIN. Dit betekent echter niet dat in Vlaanderen geen relevant onderzoek is uitgevoerd. Integendeel. Het Vlaamse onderzoek op het gebied van taal- en spraaktechnologie en aanverwante disciplines behoort al jaren tot de wereldtop. Vlaamse onderzoekers zijn de afgelopen jaren bijzonder succesvol gebleken in het binnenhalen van belangrijke financiering voor vooraanstaand onderzoek, zowel nationaal (IWT-SBO, FWO, iMinds/IMEC, BELSPO) als internationaal (CHIST ERA, EU FP7 FET, H2020 MC ITN). Dit onderzoek heeft niet alleen geleid tot talrijke wetenschappelijke publicaties, (zie bijvoorbeeld de indrukwekkende lijsten in de Appendix van dit rapport), maar heeft ook belangrijke materialen en tools opgeleverd voor verder onderzoek en ontwikkeling (*GrETEL*, *SCATE*, *Taal2Pictogram*, *SubTle*). Het feit dat deze data en tools opgenomen kunnen worden in de CLARIN-infrastructuur en dat het voor Vlaamse onderzoekers mogelijk is om gebruik te maken van materialen en tools die onderzoekers in veel andere landen bijdragen, is cruciaal voor de verdere ontwikkeling van datagedreven en technologisch gebaseerd onderzoek in de sociale en geesteswetenschappen, maar ook daarbuiten, want taal speelt een essentiële rol in vele disciplines.

Ook zonder een nationaal gefinancierd programma heeft CLARIN ERIC in Vlaanderen kunnen rekenen op een consortium van hoogstaande onderzoeksgroepen met gemotiveerde en vooraanstaande onderzoekers die op verschillende manieren met elkaar samenwerken en met collega's in het buitenland. Voorbeelden van vruchtbare internationale consortia zijn:

- het EU - Marie-Curie ITN *INSPIRE* (Investigating Speech Processing in Realistic Environments) consortium van KU Leuven, Radboud Universiteit Nijmegen, Technical University of Denmark, Philips Research Laboratories, Tampere University of Technology, Universidad del Pais Vasco, University College London, University of Edinburgh, University of York en University of Sheffield,
- het CHIST-ERA *MUSTER* (MULTimodale verwerking van ruimtelijke en temporele expressies) consortium met KU Leuven, ETH Zurich, University Pierre and Marie Curie, en University of the Basque Country,
- het H2020 - MSCA - ITN - ETN *TAPAS* (Training Network on Automatic Processing of PATHological Speech) consortium met Universiteit Gent, Idiap Research Institute, Friedrich-Alexander-Universität Erlangen Nuernberg, INESC ID - Instituto de Engenharia e Sistemas e Computadores Investigação e Desenvolvimento em Lisboa, Ludwig-Maximilians-Universität Muenchen, The Netherlands Cancer Institute - Antoni van Leeuwenhoek, Philips Research Eindhoven, Radboud Universiteit Nijmegen, Universität Augsburg, en Université Toulouse III-Paul Sabatier.

Van bijzonder belang is zeker de samenwerking met collega's in Nederland, hetgeen blijkt uit verschillende gezamenlijke projecten. Het is duidelijk dat het Vlaams-Nederlands STEVIN-programma

(Spyns & Odijk, 2013) hier een sterke impuls aan gegeven heeft. Daar werd de basis gelegd voor belangrijke samenwerkingsverbanden rondom Nederlandstalige taal- en spraaktechnologie die nu nog hun vruchten afwerpen, zoals ook blijkt uit het recent verschenen, prachtig overzicht van wat CLARIN in het Nederlands taalgebied heeft bereikt “CLARIN in the Low Countries” (Odijk & Van Hessen, 2017). In de context van CLARIN ERIC kunnen deze samenwerkingsverbanden verder worden aangevuld en uitgebouwd.

Uit dit rapport blijkt duidelijk dat Vlaamse onderzoekers bijzonder actief zijn geweest in de post-STEVIN-periode om het onderzoek dat in dat kader werd opgezet verder te laten bloeien. Het gaat zowel om fundamenteel onderzoek, denk bijvoorbeeld aan het ERC-project **COGNIMUND** (Cognitief begrijpen van beeld), aan het FP7-project **ORGANIC**, over cognitieve architectuur, of het FWO-project over datagedreven onderzoek naar ritme en prosodie in middeleeuwse teksten, en het FWO-project naar het potentieel van treebanks voor onderzoek naar Nederlands syntaxis, als om vraaggestuurd onderzoek zoals in het **ASIMOV**-project, over automatische kandidaatselectie voor vacatures, het **SubTle**-project, over preventie van zelfdoding, **ACCUMULATE**, over het detecteren van cruciale informatie in klinische teksten, en het IMEC-**TECOL**-project (Technology-Enhanced COLlaborative Learning) over de ontwikkeling, implementatie en evaluatie van recente educatief technologische oplossingen voor interactief en collaboratief leren. Het zijn projecten van grote maatschappelijke waarde over actuele onderwerpen zoals e-health, e-learning, toegankelijkheid en inclusie, en veiligheid en preventie.

Het onderzoek en de kennis die dat oplevert zijn ook van groot belang voor bedrijven. Zo heeft bijvoorbeeld UGent (LT3) de voorbije jaren een aantal consultancycontracten afgesloten in het kader van VLAIO-innovatieprojecten en haalbaarheidsstudies en hebben bedrijven als Hello Customer, TheMatchBox, The Vigor Unit en EBMPracticeNet gebruik kunnen maken van hun kennis en expertise. De technologie die ontwikkeld is door UAntwerpen en KU Leuven heeft ook al de weg gevonden naar commerciële toepassingen via spin-off-bedrijven zoals Textgain (<https://www.textgain.com/>) en Fluent AI (<https://www.fluent.ai/>).

De toekomst ziet er dan ook rooskleurig uit voor het taal- en spraaktechnologie-onderzoek in Vlaanderen, ook in tijden van FAMGA (Facebook, Apple, Microsoft, Google, and Amazon), zoals onlangs werd geconstateerd ter gelegenheid van het 30-jarig bestaan van de spraaktechnologie-groep ESAT van KU Leuven. Ondanks het feit dat de grote bedrijven steeds dominanter worden in het ontwikkelen van de kerntechnologie, kunnen universiteiten nog veel bijdragen als het gaat om onderzoek naar het inzetten van die technologie in realistische maatschappelijke toepassingen en voor kleinere talen. En omdat data, corpora en infrastructuur steeds belangrijker worden en de dataverzamelingen steeds groter moeten zijn, is aansluiting bij grootschalige initiatieven die onderzoekers kunnen ondersteunen en voorzien van relevante data en tools om die data te kunnen gebruiken, langzaam maar zeker onmisbaar aan het worden. Met andere woorden, de Vlaamse deelname aan de CLARIN ERIC is nu actueler en noodzakelijker dan ooit, zoals blijkt uit de rest van dit rapport.

De structuur van dit rapport is als volgt. In het volgende hoofdstuk wordt goed uitgelegd wat de voordelen zijn van CLARIN ERIC, niet alleen voor Vlaamse onderzoekers, maar ook voor Vlaanderen zelf. Vervolgens geven we een beknopt overzicht van de samenwerking binnen het CLARIN-ERIC-Vlaanderen-consortium in de periode 2012-2017. Daarop volgt een overzicht van onderzoeks- en

ontwikkelingsprojecten op het gebied van taal-en spraaktechnologie en aanverwante disciplines die zijn uitgevoerd aan Vlaamse universiteiten in dezelfde periode. In het volgende hoofdstuk worden enkele voorbeelden gegeven van strategische netwerken en samenwerkingsverbanden ten behoeve van valorisatie en maatschappelijke benutting. In het afsluitende hoofdstuk schetsen we de perspectieven voor de toekomst.

CLARIN: EEN MEERWAARDE VOOR VLAAMSE ONDERZOEKERS EN VLAANDEREN

Het hoofddoel van CLARIN is het opzetten en onderhouden van een veilige en duurzame infrastructuur voor taalmaterialen, zodat die eenvoudig gedeeld, (her)gebruikt, geanalyseerd en/of verrijkt kunnen worden. Deel uitmaken van het CLARIN-netwerk levert een aantal voordelen op, zowel voor individuele Vlaamse onderzoekers, als voor Vlaanderen zelf.

A. VOORDELEN VOOR ONDERZOEKERS

i. RAADPLEGEN VAN DATA

In CLARIN staat de vindbaarheid en/of de raadpleging van taalmaterialen centraal. Onderzoekers kunnen met behulp van metadata materialen zoeken in de CLARIN-catalogus, de Virtual Language Observatory (VLO). De VLO bevat verwijzingen naar meer dan 1,6 miljoen materialen in zowel nationale als regionale talen. Er kan ook in een aantal dataverzamelingen zelf gezocht worden, met behulp van de Content Search. Waar mogelijk behandelt CLARIN de data in overeenstemming met de FAIR-principes die in het kader van de aandacht voor Open Science zijn ontwikkeld: Findability, Accessibility Interoperability, Re-usability.

Onderzoekers kunnen met behulp van de VLO snel en eenvoudig relevante taalmaterialen vinden die ze kunnen hergebruiken of uitbreiden. Bovendien biedt de VLO de mogelijkheid aan om met (delen van) bestaande dataverzamelingen eigen datasets te creëren. Wanneer de metadata daarvan gepubliceerd wordt in de VLO, worden die “nieuwe” datasets beter zichtbaar voor anderen en kunnen ze gebruikt worden om onderzoek te repliceren.

ii. ARCHIVEREN/CITEREN VAN DATA

De CLARIN-infrastructuur streeft ernaar om taalmaterialen te archiveren en beschikbaar te stellen van de onderzoeksgemeenschap en dat op een veilige en duurzame manier. Per CLARIN-lid is er minstens 1 CLARIN-center dat onderzoekers helpt om hun materialen (e.g. corpora, lexica, audio- en video-opnames, annotaties, grammatica's, enz.) op te slaan en beschikbaar te stellen van anderen. De voordelen hiervan zijn:

- Het gaat om archivering op lange termijn.
- Andere onderzoekers kunnen materialen hergebruiken, verbeteren of erop voortbouwen.
- De CLARIN-centers geven advies op gebied van IPR en licenties en nemen o.a. de gebruikerslicentieafhandeling op zich.

- Het is mogelijk om materialen eenvoudig en betrouwbaar te citeren: datasets en tools krijgen een “persistent identifier” in de CLARIN-infrastructuur, zodat er niet langer geciteerd moet worden m.b.v. instabiele URL’s.

iii. BEWERKEN VAN DATA

CLARIN biedt uitstekende tools en online diensten aan voor vele talen. Taalmaterialen kunnen gecombineerd, geanalyseerd en verrijkt worden met behulp van een aantal linguïstische tools die beschikbaar zijn binnen de CLARIN-infrastructuur (bv. automatische POS-tagging, informatie-extractie, audio-/video-analyse...). Het automatisch annoteren van grote hoeveelheden data kan eenvoudiger en sneller uitgevoerd worden en data en tools van verschillende centra kunnen gecombineerd worden. Voor zeer grote en computationeel intensieve taken, kan CLARIN onderzoekers verbinden met vooraanstaande High Performance Computing (HPC) centers.

iv. EENMALIGE AANMELDING

CLARIN heeft een Service Provider Federation tot stand gebracht. Dat is een betrouwbaar netwerk van identity providers die een “eenmalige aanmelding” aanbieden. Dat houdt in dat onderzoekers in kunnen loggen met de toegangsgegevens van hun eigen instelling om toegang te krijgen tot afgeschermdedataverzamelingen en tools in andere instellingen en landen. Op die manier winnen ze tijd en hoeven ze slechts 1 wachtwoord bij te houden voor de gehele CLARIN-infrastructuur.

v. TOEGANG TOT KENNIS

CLARIN biedt niet enkel toegang tot data en tools, maar ook tot kennis en expertise, via de zogenaamde Knowledge Sharing Infrastructure (KSI).

- Elk CLARIN-center dat toegang tot data en tools aanbiedt, bemant ook een helpdesk waar gebruikers informatie kunnen krijgen over (het gebruik van) die materialen of bv. over het ontwikkelen van taalmaterialen. Daarnaast bemant het hoofdkantoor van CLARIN ERIC een centrale helpdesk, die indien nodig nationale experts inschakelt.
- Gecertificeerde CLARIN-kenniscentra bieden toegang tot expertise in specifieke domeinen, zoals bv. treebanking, spraakanalyse, de Poolse taal,...
- Speciale commissies binnen CLARIN verzamelen kennis over bv. IPR en licenties (CLIC – CLARIN Legal Issues Committee) en standaarden (STAC - Standards Committee).
- Een belangrijk luik van de KSI bestaat uit CLARIN-bijeenkomsten: bv. workshops omtrent specifieke thema’s of het CLARIN Annual Congress, waar infrastructuraanbieders en gebruikers uit alle CLARIN-landen expertise en ideeën uitwisselen.
- Er kan een subsidie (een CLARIN mobility grant) aangevraagd worden voor grensoverschrijdende, korte verblijven in CLARIN-centra die de uitwisseling van kennis en expertise bevorderen tussen de werknemers van verschillende centra of tussen de werknemers van een center en onderzoekers.

- De KSI biedt ook online cursussen aan en presenteert use cases, zodat het eenvoudiger wordt voor onderzoekers om nieuwe technieken onder de knie te krijgen, of zodat het duidelijk wordt welke tools/materialen bruikbaar zijn voor bepaalde onderzoeksvragen.

vi. CENTRALE DIENSTEN

Onderzoekers uit de CLARIN-community kunnen gebruikmaken van een uitgebreid gamma aan centrale CLARIN-diensten, die ontwikkeld worden in een nauwe samenwerking tussen CLARIN-ERIC en de nationale teams. Het gaat dan onder meer om de centrale Discovery Service (<https://www.clarin.eu/content/clarin-discovery-service>), diensten voor het creëren, delen en hergebruiken van uitgebreide metadatabeschrijvingen, hulp bij het creëren van datamanagementplannen.

vii. DEELNAME IN EUROPESE PROJECTEN

CLARIN streeft naar deelname in Europese projecten wanneer die de algemene strategie en/of de verdere ontwikkeling van de CLARIN-infrastructuur ondersteunen. Dit houdt in dat CLARIN-leden een verhoogde kans hebben op deelname in door Europa gesubsidieerde projecten aangezien de werknemers van nationale consortia (of hun vertegenwoordigers) aan zulke projecten mogen werken in naam van CLARIN ERIC. Het gaat/ging bijvoorbeeld om de volgende projecten: het H2020 project CLARIN-PLUS, PARTHENOS, EUDAT2020, Language Technology Observatory en EUROPEANA-DSI en EOSC-hub.

B. VOORDELEN VOOR VLAANDEREN

Een groot voordeel van CLARIN is de schaalvergroting: algemene infrastructuurdiensten kunnen over grenzen heen gebruikt worden, waardoor de kost van het opzetten en onderhouden van die structuren gedeeld kan worden tussen landen. De toegang tot de enorme rijkheid aan CLARIN-materialen zelf kan ook leiden tot meer geavanceerd onderzoek, tot dalende ontwikkelingskosten en tot nieuwe samenwerkingsmogelijkheden over grenzen en vakgebieden heen, wanneer onderzoekers van over heel Europa verder kunnen bouwen op elkaars resultaten.

Het CLARIN-lidmaatschap draagt bij aan een betere positie van het Nederlands/Vlaams in een Europese (en internationale) context doordat de Nederlandse/Vlaamse data meer zichtbaar en eenvoudiger toegankelijk worden voor een breder publiek, net zoals het culturele erfgoed dat inherent deel uitmaakt van taalmaterialen in de CLARIN-infrastructuur.

Het lidmaatschap van CLARIN vergroot ook het netwerk van de Vlaamse onderzoekers: het beëindigen van het lidmaatschap van Vlaanderen aan CLARIN staat gelijk aan het onttrekken van de Vlaamse onderzoekers aan een belangrijk Europees netwerk, dat steeds groter wordt en nu al 19 leden telt (en 2 waarnemers).

Het is ook van belang dat CLARIN-leden een directe invloed hebben op het beslissingsproces i.v.m. alle aspecten van de infrastructuur (de constructie en het opereren ervan, de evolutie op lange termijn, strategische prioriteiten,...), aangezien de CLARIN-leden vertegenwoordigd zijn en kunnen

meestemmen in de General Assembly, het National Coordinators' Forum, etc., en door deelname in verscheidene commissies.

En ten slotte is het ook belangrijk om te benadrukken dat hoewel de CLARIN ERIC-infrastructuur in eerste instantie ontwikkeld is voor onderzoekers, het ook voor bedrijven relevant is. Binnen CLARIN zijn intussen heel veel materialen en tools ontwikkeld in vele verschillende talen. Meer dan de helft daarvan is vrij toegankelijk en kan dus ook door bedrijven worden gebruikt. Het CLARIN-consortium is ook van plan om hands-on workshops te gaan organiseren voor bedrijven om het gebruik van CLARIN-resources te faciliteren.

CLARIN ERIC VLAANDEREN 2012-2017

Eind 2011 werd er een overeenkomst getekend tussen het Algemeen Secretariaat (AS) van de Taalunie en de Vlaamse regering waarin de Vlaamse regering zich ertoe verbond het AS te subsidiëren voor zijn deelname aan de CLARIN ERIC als vertegenwoordiger van Vlaanderen en voor het opzetten van CLARIN-activiteiten in Vlaanderen. De uitwerking van het besluit werd opgevolgd door EWI.

Het AS heeft in 2012 als stichtend lid meegewerkt aan de oprichting van de CLARIN ERIC, de concrete CLARIN-activiteiten werden hoofdzakelijk uitgevoerd door het Instituut voor Nederlandse Lexicologie (INL) te Leiden, meer bepaald door de TST-Centrale. Het doel was om de CLARIN-infrastructuur en (de materialen van) de TST-Centrale bekendheid te geven in Vlaanderen door middel van disseminatie-, netwerk- en trainingsactiviteiten.

De verhuizing van de TST-Centrale (in een verminderde capaciteit) naar de Nederlandse Taalunie zorgde begin 2013 ervoor dat de CLARIN-werkzaamheden moeizaam verliepen, omdat de technische (CLARIN-)infrastructuur en (CLARIN-) personeel bij het INL bleven en o.a. omdat taakverdeling tussen TST-Centrale/Taalunie en het INL onduidelijk was.

De beginperiode was dus een uitdaging, maar in de daaropvolgende jaren werd er langzaam maar zeker verder gebouwd aan een CLARIN in Vlaanderen en hieronder worden in grote lijnen een aantal relevante verwezenlijkingen uit die periode aangehaald. Voor meer details verwijzen wij u graag door naar de eerder opgeleverde CLARIN-EWI-NTU-jaarverslagen.

Het INL, het AS en vertegenwoordigers uit het Vlaamse veld maakten van bij het begin deel uit van de CLARIN-overlegorganen/-werkgroepen: het National Coordinators forum, het CLARIN Legal (and Ethical) Issues Committee, het User Involvement Committee, het Standing Committee for CLARIN Technical Centres, de General Assembly en het Programme Committee van de CLARIN Annual Conference. Het consortium was bovendien ook actief betrokken bij de ontwikkeling van CMDI (de Component MetaData Initiative) en van de CLARIN Concept Registry.

Ook op andere manieren werd de relatie met CLARIN verankerd. Zo heeft het INL een CLARIN-centre opgericht, dat zich richtte op Vlaanderen en Nederland. De TST-Centrale en dat centre werkten samen t.b.v. de CLARIN-kwaliteitscontrole en in 2014 heeft het INL CLARIN Centre het Data Seal of Approval gekregen (wat verlengd werd in 2017). De materialen van het INL en van de TST-Centrale werden ook geïntegreerd in de CLARIN-infrastructuur. Enerzijds door opname van de metadata in Virtual Language Observatory, zodat de materialen vindbaar zijn voor de gehele CLARIN-community en anderzijds doordat het INL deel ging uitmaken van de CLARIN Service Provider Federation (een betrouwbaar netwerk van identity providers die een “eenmalige aanmelding” aanbieden). Dat heeft als gevolg dat onderzoekers in kunnen loggen met de toegangsgegevens van hun eigen instelling om toegang te krijgen tot afgeschermd INL-dataverzamelingen en tools.

Er is gewerkt aan de promotie van CLARIN en TST bij Vlaamse ondernemingen en onderzoekers via verscheidene communicatiekanalen zoals de TST-nieuwsbrief, Twitter, de website/webwinkel. Bovendien werden er ook regelmatig congressen en netwerkevenementen bezocht door mensen van het AS, het INL of door vertegenwoordigers uit het Vlaamse veld, met als doel het hergebruik van taalmaterialen te stimuleren en om kennis in verband met TST en CLARIN te delen. Er werden eveneens Nederlandse en Vlaamse (MKB-/KMO-)bedrijven opgezocht om in gesprek te gaan over de mogelijkheden die TST en TST-taalmaterialen boden.

Het CLARIN-Vlaanderen-consortium heeft een aantal evenementen zelf georganiseerd. In 2013 werd het symposium "E-health voor mensen met communicatieve beperkingen: wetenschap en praktijk" georganiseerd op de Hogeschool van Arnhem en Nijmegen. In een afwisselend programma werden wetenschap en praktijk uit Nederland en Vlaanderen bij elkaar gebracht. Deskundigen uit de praktijk, onderzoekers en taal- en spraaktechnologen presenteerden geavanceerde mogelijkheden om e-health-toepassingen in te zetten voor de behandeling en ondersteuning van mensen met een communicatieve beperking en bespraken de voordelen die therapeuten en cliënten daarbij ondervinden. In 2015 werd een copy-editingbijeenkomst in Antwerpen georganiseerd ter ondersteuning van Nederlandstalige publicatietrajecten in de Nederlandse en Vlaamse uitgeverwereld, Taal in Bedrijf 2015 in Brussel, en de 24^{ste} editie van CLIN in Leiden. In 2016 werd in Leuven de CLARIN-PLUS-workshop "Working with Digital Collections of Newspapers" georganiseerd en colleges voor studenten taalkunde in Antwerpen en Leiden i.v.m. het Corpus Gesproken Nederlands en de CGN-exploratiesoftware Corex. Ook makel- en schakelactiviteiten waren belangrijk. Er werd bijvoorbeeld een nieuw, permanent onderdeel van het MuZIEum in Nijmegen mogelijk gemaakt, waar getoond wordt wat Nederlandstalige TST kan betekenen voor blinden en slechtzienden. Bedrijven (waaronder het Vlaamse Sensotec) kunnen daar hun toepassingen tentoonstellen aan de vele bezoekers waaronder ook veel scholieren zitten, hetgeen ook een educatieve waarde heeft.

Het Algemeen Secretariaat bood (Vlaamse) onderzoekers de mogelijkheid om deel te nemen aan (CLARIN-)netwerkbijeenkomsten door het verstrekken van subsidies voor de reis- en verblijfskosten of deelnamekosten, een mogelijkheid waarvan meermaals gebruik is gemaakt. De servicedesks van de TST-Centrale en het INL waren eveneens essentiële kanalen bij het verspreiden van informatie en kennis i.v.m. TST en CLARIN, maar ook i.v.m. IPR of het opbouwen van een taalmateriaal. Naast een informatieve rol hadden ze ook een dienstverlenende functie: onderzoekers met concrete onderzoeksvragen werden begeleid tijdens hun zoektocht naar geschikte materialen en tools in functie van hun onderzoek, bedrijven die data(-op-maat) nodig hadden voor onderzoek en ontwikkeling konden (tegen betaling) een beroep doen op de servicedesks,...

De beschikbare taalmaterialen bekendheid geven was een eerste doel, maar het was ook essentieel om het hergebruik ervan zo veel mogelijk te stimuleren. Het ontwikkelen van een overzichtelijke webcatalogus met online bestelmodule droeg daaraan bij, net zoals het feit dat ernaar gestreefd werd om alle materialen online toegankelijk of downloadbaar te maken, vergezeld van alle relevante documentatie. Ook commercieel gebruik van dataverzamelingen was mogelijk. Bij dit aspect van de terbeschikkingstelling werd de TST-Centrale bijgestaan door een Prijzencommissie, bestaande uit vertegenwoordigers uit het Vlaamse en Nederlandse

onderzoeksveld en de Vlaamse en Nederlandse bedrijfswereld. In eerste instantie had de commissie enkel de taak om advies te geven in verband met commerciële beschikbaarstelling voor materialen uit de TSTC-catalogus, maar na verloop van tijd kreeg ze een verruimde taakstelling en kon ze zich ook uitspreken over strategische kwesties op het vlak van inrichting, organisatie, beheer, enz. Dat resulteerde in het besluit om soepeler om te springen met de gebruikerslicenties, waardoor het hergebruik van de materialen voorop stond. Zo is toegestaan dat corpora onder dezelfde geldende licentievooraarden digitaal beschikbaar worden gesteld aan de gebruikers van Nederlab, ook aan die uit Vlaanderen en aan de nieuwe automatische vertaalmachine CEF.AT van de Europese Unie.

In 2016 kwam er een einde aan de overeenkomst tussen het AS van de Taalunie en de Vlaamse regering. Ondertussen had het Instituut voor de Nederlandse Lexicologie zich ongevormd tot een breder Instituut voor de Nederlandse Taal (INT) en was de TST-Centrale met haar materialen overgedragen aan dat INT, waardoor de materialen een plaats binnen het INT CLARIN Centre kregen.

Het Algemeen Secretariaat en het INT hebben de CLARIN-activiteiten op eigen initiatief en op eigen kosten verdergezet en de rol als CLARIN-vertegenwoordiger van Vlaanderen werd uitgebouwd.

In eerste instantie heeft het INT in Vlaanderen een bevraging uitgevoerd om de huidige benutting van CLARIN door Vlaamse onderzoekers zichtbaar te maken. Vervolgens werd er in de herfst van 2016 een gespreksronde in Vlaanderen opgestart om naar verwachtingen van Vlaamse onderzoekers ten aanzien van het INT zelf te peilen, maar ook om de mogelijkheden van TST en CLARIN onder de aandacht te brengen.

Tijdens die Tour de Flandre ging de INT-directeur Frieda Steurs het gesprek aan met verschillende organisaties en personen in Gent, Leuven en Antwerpen en werd er ook doorlopend overlegd met Franciska de Jong, de CLARIN-directeur en Jan Odijk (CLARIAH-NL¹). Er vonden ook gesprekken met DARIAH² Vlaanderen plaats om na te gaan welke mogelijkheden er waren om samen te werken.

Uit de gesprekken en de bevraging bleek dat er een paar belangrijke taken weggelegd zijn voor het INT CLARIN Centre, zoals expertise en ondersteuning aanbieden aan onderzoekers die geen of weinig ervaring hebben met digitaal werken, bv. door hands-on workshops te organiseren. Een andere taak is dat het INT ervoor kan zorgen dat aan het einde van een project de opgebouwde expertise en taalmaterialen niet verloren gaan en bovendien ter beschikking gesteld worden van anderen (via de CLARIN-infrastructuur). Er is ook gebleken dat DARIAH-Vlaanderen positief staat t.o.v. een intensieve samenwerking met CLARIN teneinde de krachten te kunnen bundelen.

Er werd eveneens gewerkt aan de bekendheid van het INT (als CLARIN-centre) en van CLARIN door het geven van verschillende presentaties op evenementen in Vlaanderen en Nederland (Leuven, Antwerpen, Utrecht, Leiden, Brussel,...) en in oktober 2017 gaf het INT een eerste CLARIN-workshop in Vlaanderen voor de digital humanists aan de Universiteit van Antwerpen. Op

¹ <https://www.clariah.nl/>

² DARIAH is een internationaal consortium dat gericht is op het verbeteren en ondersteunen van digitaal onderzoek en onderwijs binnen de geesteswetenschappen. Zie ook <http://www.dariah.eu> en <http://be.dariah.eu/dariah-vl>.

deze studienamiddag presenteerde het INT zich en werd er nader ingegaan op de meerwaarde van het INT en CLARIN voor de digital humanities op het gebied van tools.

Het hergebruik van materialen bleef steeds centraal staan en teneinde het commerciële gebruik van de voormalige TSTC-materialen optimaal te stimuleren, werd er eind 2017 beslist dat commerciële partijen (de meeste van) die materialen kosteloos kunnen gebruiken voor onderzoek en ontwikkeling. Over die beslissing zal in de nabije toekomst uitgebreid gecommuniceerd worden na de lancering van de vernieuwde en geïntegreerde INT-webwinkel, waarin alle materialen van het INT zelf en de materialen van de voormalige TSTC-Centrale geïntegreerd zullen worden.

Op 12 oktober 2017 werd op het INT ook een bijeenkomst georganiseerd voor bedrijven uit Nederland en Vlaanderen om hen te informeren over de materialen die in de loop der jaren via CLARIN-ERIC beschikbaar zijn gekomen en die voor hen toegankelijk zijn.

De afgelopen periode bleef Vlaanderen vertegenwoordigd binnen CLARIN door een actieve deelname aan de CLARIN-overlegorganen en –werkgroepen die hierboven reeds vermeld werden en werd er ook expertise uitgewisseld met andere CLARIN-partners. Zo werkte de teamleider softwareontwikkeling van het INT bv. mee aan een miniworkshop over Korp, Strix en BlackLab bij [Språkbanken](#) (de Zweedse Taalbank) in Gothenburg, met steun van CLARIN in de vorm van een CLARIN Mobility Grant.

Ten slotte is in de periode februari maart 2018 het Vlaamse CLARIN-ERIC-consortium in de focus van de Tour de CLARIN: <https://www.clarin.eu/blog/tour-de-clarin-dluflanders>.

ONDERZOEK TAAL- EN SPRAAKTECHNOLOGIE EN DIGITAL HUMANITIES IN VLAANDEREN 2012-2017

A. KU LEUVEN – OVERZICHT PROJECTEN

1. Titel: STON: Spraak- en Taaltechnologisch Ondertitelen in het Nederlands (2014-2016)

Financiering: IWT O&O

Partners:

- a. KU Leuven (ESAT) (BE)
- b. Devoteam (BE)
- c. Limecraft (BE)
- d. Pervoice (IT)
- e. UGent (ELIS – DSL)(BE)

Beschrijving: STON is een innovatief aanbestedingsproject dat via spraak- en taaltechnologie ondersteuning biedt bij het ondertitelen van tv-programma's. Er worden verschillende componenten ontwikkeld: taalidentificatie en sprekerdiarisatie, spraakherkenning, postprocessing om de transcriptie om te vormen in een goed geformatteerde ondertitel. Er wordt ook een gebruikersinterface ontwikkeld die de componenten aanstuurt, en de ondertitelaar toelaat om efficiënt het resultaat te verifiëren en aan te passen indien nodig. De focus in het project ligt op het automatisch ondertitelen van documentaires en online nieuwscontent.

Output: tools: spraak/niet-spraak segmentatie, taalidentificatie, sprekerdiarisatie, aangepaste akoestische en taalmodellen voor spraakherkenning van het Vlaams

Outreach: VRT Sandbox, voorstelling aan het Vlaams Parlement

2. Titel: ALADIN: Adaptation and Learning for Assistive Domestic Vocal Interfaces (2011-2014)

Financiering: IWT – SBO

Partners:

- a. KU Leuven (ESAT)
- b. UAntwerpen (CLiPS)
- c. KU Leuven (Advise - (bij aanvang: MOBILAB))
- d. KU Leuven (MintLab (bij aanvang: CUO))
- e. IN-HAM (subcontractor)

Beschrijving: ALADIN is een taalafhankelijke spraakinterface, bedoeld voor mensen met een motorische beperking of verlamming. Vanwege hun handicap is het gebruik van traditionele afstandsbedieningen moeilijk of onmogelijk. Vele gerelateerde aandoeningen veroorzaken ook spraakstoornissen, wat traditionele spraakherkenningssystemen ongeschikt maakt voor deze doelgroep. ALADIN is in staat om woordenschat te leren en grammatica te induceren op een robuuste en adaptieve manier. Gebruikers hoeven hun taalgebruik niet aan te passen aan het systeem en kunnen hun eigen woorden en zinnen gebruiken om bijvoorbeeld hun huis en apparaten te bedienen.

Output: Software in licentie aan Fluent.AI, een startup die nu 15 personeelsleden telt. Databanken met dysartrische spraak. Doctoraat Joris Driesen en Bart Ons.

Outreach: Slotsymposium voor het werkveld in "Care Home of the Future", Heusden-Zolder, 9 dec 2014

3. Titel: SPRAAK2TAAL: extractie van geladenheid en ideologische context uit politieke redevoeringen - een onderzoek naar integratie van spraak- & taalverwerking (2010-2012)

Financiering: EWI

Partners:

- a. KU Leuven (ESAT)
- b. KU Leuven (LIIR)

Beschrijving: Dit project onderzoekt hoe spraakherkenning en een taalmodule optimaal kunnen samenwerken in het CLARIN-raamwerk. Als case-studie is gekozen voor "extractie van geladenheid en ideologische context uit politieke redevoeringen". Het project (1) onderzoekt het nut van de verschillende types uitvoer die de spraakherkenner kan genereren (beste zin, zin met parallelle alternatieven, of woordgrafe) voor de taalmodule, (2) past daartoe de uitvoertypes van de SPRAAK-herkenningssoftware aan, en (3) past automatisch het taalmodel van de herkenner aan aan het type taalgebruik (politieke redevoeringen) d.m.v. probabilistische latente modellen.

Output: verbeteringen en uitbreidingen aan de spraakherkenningstools

4. Titel: AMODA: Acoustic modeling for under-resourced languages using feature space constraints - applied to the South African language situation (2010-2015)

Financiering: FWO

Partners:

- a. KU Leuven (ESAT)
- b. CSIR (Zuid-Afrika)

Beschrijving: State-of-the-art spraakherkenningssystemen bevatten miljoenen parameters die getraind worden aan de hand van voorbeeldgegevens. Deze parameters vertonen een grote redundantie, en men vertrouwt op enorme corpora (honderden uren) voor het trainen ervan. Hoewel er zeer grote corpora beschikbaar zijn voor grote talen, is dit niet het geval voor kleinere talen. Het doel van dit project is om nieuwe wiskundige technieken toe te passen die de redundanties in de parameters uitbuiten en toegepast kunnen worden op talen met een tekort aan resources. Bij alle technieken is de basisveronderstelling dat een globale beschrijving van de akoestische ruimte van spraak zeer gelijkend is van taal tot taal en dat deze component gedeeld kan worden met andere talen (of slechts licht aangepast moet worden). Anderzijds moet de categorische indeling per klank taal per taal aangeleerd worden. In dit project werd enerzijds gewerkt met (i) subspace technieken en (ii) diepe neurale netwerken. Verschillende bron/target-scenario's werden uitgetest: zowel Afrikaans (low resource target) op basis van Nederlands (brontaal) als een complexere set-up met vele bron- en targettalen.

Output: Doctoraat van Reza Sahraeian (gedeeltelijk)

Outreach: In het kader van de samenwerking met Zuid-Afrika werd een winterschool (1 week) over spraakherkenning georganiseerd in Pretoria in 2014 en later in 2016 nogmaals in Stellenbosch. Beide events stonden open voor onderzoekers van over heel Zuid-Afrika en de buurlanden.

5. Titel: SCATE: Smart Computer-Aided Translation Environment (2014-2018)

Financiering: IWT-SBO

Partners:

- a. KU Leuven (CCL) - coördinator
- b. KU Leuven (ESAT)
- c. KU Leuven (LIIR)
- d. KU Leuven (QLVL)
- e. UGent (LT3)

f. Universiteit Hasselt (EDM)

Beschrijving: In het SCATE-project werden verschillende aspecten van vertaaltechnologie onderzocht en het gebruik daarvan binnen de workflow van professionele vertalers, met als doel het vertaalproces te optimaliseren. Er werd in nauwe samenwerking met de vertaalindustrie onderzoek gedaan naar

- automatische vertaling (bv. met behulp van syntactic fuzzy matching) en hoe deze te integreren met vertaalgeheugens
- automatische evaluatie van MT-output, zodat ingeschat kan worden wanneer deze te gebruiken is
- automatische terminologie-extractie uit parallelle en vergelijkbare corpora
- spraakherkenning in de context van MT, zoals spraakvertaling en spraakherkenning als input-methode voor post-editing
- effect van verklaarbaarheid en begrip (intelligibility) van de informatiebronnen in de interface op post-editie

Daarnaast werd onderzocht hoe de gebruikersinterface van de vertaler optimaal ontworpen kon, wat leidde tot een prototype-demonstratiesysteem.

Output:

- SCATE-prototype (een vertaalomgeving) + onderliggende modules
- GrETEL 3.0
- doctoraten van Joris Pelemans (gedeeltelijk), Lyan Verwimp (nog af te leggen, gedeeltelijk) en Tom Vanallemeersch (gedeeltelijk), Arda Tezcan (14 maart 2018), Geert Heyman (verwacht in 2018)
- verbeteringen en aanpassingen aan de spraakherkenningstools
- Poly-GrETEL (o.a. voor parallelle corpora)
- LENG (Lexically Equivalent Node Grouping - voor alignering)
- SeRLino (Semantic Role Linking through nodes - voor alignering)
- Bilingual Lexicon induction
- TM-MT-integratiemethode

Outreach:

- Springer EAMT-2015 Best Paper Award
- vertalers, post-editors, vertaalindustrie, opleidingen tolk-vertaler, taalkundigen
- Language Industry Awards (plaats 9 in Top 100)
- VLAIO O&O met Contact.Fit
- SCATE werd gepresenteerd op volgende conferenties voor professionele vertalers of een breed publiek:
 - o International Annual Meeting on Computer-Assisted Translation and Terminology JIAMCATT 2017, Den Haag
 - o Translating and the Computer, 2015 & 2017, London
 - o Tekom, 2016 Stuttgart
 - o The future of language technology – a conference for translation staff (2016) European Economic and Social Committee and Committee of the Regions, Brussel

- o LT-Innovate (2016, 2017)
- o European Symposium on Languages for Special Purposes 2017
- o NLP Meetup Leuven (2016)
- o Annual Meetings of the European Association for Machine Translation (2014-2017)
- o TAO-CAT 2015

6. Titel: TTNWW: TST Tools voor het Nederlands als Webservices in een Workflow (2010-2012)

Financiering: EWI (voor het Vlaamse gedeelte)

Partners:

- a. KU Leuven (CCL) - coördinator
- b. KU Leuven (ESAT)
- c. UGent
- d. UAntwerpen (CLiPS)
- e. KADOC (Leuven)
- f. Radboud Universiteit (Nijmegen, NL)
- g. Universiteit Twente (Enschede, NL)
- h. Aletta (Amsterdam, NL)
- i. Meertens Instituut (Amsterdam, NL)

Beschrijving: Het doel van dit project is om allerlei bestaande componenten die ontwikkeld zijn in (o.a.) CGN en STEVIN in te passen in een workflowsysteem voor web services dat (o.a. met grote Nederlandse inbreng, bijv. MPI) ontwikkeld wordt in CLARIN-verband, en dit geheel te laten draaien op servers van erkende CLARIN-centra, met als doel faciliteiten aan te bieden voor onderzoekers uit de HSS met geen of weinig technische bagage. Deze faciliteiten moeten 1) hen in staat stellen hun onderzoeksvragen beter of makkelijker aan te pakken en 2) mogelijkheden bieden voor het formuleren van nieuwe typen onderzoeksvragen, i.e. onderzoeksvragen die voor CLARIN niet gesteld konden worden of niet doelmatig te beantwoorden waren. De webservices betreffen twee modaliteiten: tekst en spraak.

Output: webservices die toegang geven tot de ontwikkelde tools

7. Titel: ASSIST: Assistive Speech Interface for Smart Technologies (2016-2019)

Financiering: FWO Strategische Onderzoeksbeurs

Partners: KU Leuven (ESAT)

Beschrijving: In het ASSIST-project wordt een spraakinterface voor commando- en besturingstoepassingen ontworpen. Gesproken commando's van de gebruiker en de

bijbehorende taken worden geleerd door interactie met die gebruiker te demonstreren met een alternatieve interface. Omdat ASSIST leert van interactie met de gebruiker, is het perfect op hem of haar afgestemd. De gebruiker kan zijn of haar eigen frasering van de commando's kiezen. ASSIST is taalafhankelijk, zodat het zonder aanpassingen overal ter wereld kan worden gebruikt. Een dialect of pathologische spraak is geen fundamenteel probleem meer, en er zijn dus veel toepassingen mogelijk voor mensen met een handicap.

Output:

- databank met robotcommando's
- software: TF-kaldi die toelaat om modellen voor spraakherkenning te gebruiken in de open source spraakherkenningssoftware "Kaldi"
- Nabu: software voor het trainen en testen van end-to-end ANN-spraakherkenners.
- doctoraat Vincent Renkens (nog af te leggen)

Outreach:

- presentatie op LICT-event "deep learning", 2017
- pitch en demo op DSP-Valley event "Smart Systems Industry Summit", 2017

8. Titel: CAMETRON: Towards a Virtual Director (2014-2018)

Financiering: KU Leuven – GOA

Partners: KU Leuven (ESAT)

Beschrijving: Binnen CAMETRON streven we naar een systeem dat met minimale menselijke tussenkomst audiovisuele (AV) producties van hoge kwaliteit produceert. Dit wordt bereikt door het gebruik van de nieuwste stand van zaken op het gebied van verschillende onderliggende technologieën, zoals AV-inhoudsanalyse en interpretatie van de omgeving (met focus op mensen), gezamenlijke verwerking van meerdere audio- en videostromen, multimodale integratie, realtime verwerking, (pro-)actieve besturing van camera's of microfoons, passende coderingsschema's, quadcopterbesturing, enz.

Output: doctoraat Sayeh Mirzaei

9. Titel: BBFOR2: Bayesian Biometrics for Forensics (2010-2014)

Financiering: EU - Marie Curie ITN

Partners:

- a. KU Leuven (ESAT), BE
- b. Stichting Katholieke Universiteit Nijmegen, NL

- c. Universidad Autonoma de Madrid, ES
- d. Politecnico di Torino Polito, IT
- e. Universiteit Twente, UTW NL
- f. IDIAP (Fondation de l'Institut Dalle Molle d'Intelligence Artificielle Perceptive), CH
- g. University of York, UK
- h. Höskolan i Halmstad, SE
- i. Netherlands Forensic Institute, NL
- j. Agnitio Voice Biometrics Agnitio, ES
- k. Netherlands organisation for applied research (TNO), NL
- l. Forensic Science Service, UK

Beschrijving: In dit project wordt onderzoek verricht voor biometrische detectie van spreker-, vingerafdruk- en gezichtsherkenning. Het doel is de ontwikkeling van een Bayesiaans kader voor het presenteren van multimodaal biometrisch bewijsmateriaal. ESAT heeft zich geconcentreerd op sprekerprofilering, d.w.z. het afleiden van metadata zoals geslacht, leeftijd en grootte op basis van stemmateriaal.

Output: doctoraat Hasan Bahari

10. Titel: BATS: Topic and Speaker Tracking in Broadcast Archives (2008-2012)

Financiering: IBBT, ICTRegie

Partners:

- a. KU Leuven (ESAT) (BE)
- b. Radboud Universiteit (CLST) (NL)
- c. VRT (BE)

Beschrijving: In dit project is de kennis en technologie ontwikkeld om snel, gericht en flexibel te kunnen zoeken naar en in multimediodocumenten in de archieven van de radio- en tv-stations in Vlaanderen en Nederland. Met name wordt het mogelijk om (fragmenten van) documenten te vinden die over specifieke onderwerpen gaan of waarin een specifieke persoon spreekt. Voor beide doelen is het nodig om opgenomen programma's (nieuws, actualiteiten, documentaires, etc.) te segmenteren in fragmenten die over één onderwerp handelen en/of in fragmenten waarin één spreker aan het woord is. De opnamen moeten ook gesegmenteerd worden in fragmenten waarin wel, c.q. niet gesproken wordt, en alle spraak moet automatisch getranscribeerd worden zodat de fragmenten geïndexeerd kunnen worden.

Output: doctoraat Xueru Zhang

11. Titel: VASI: Vocabulary Acquisition from Speech and Images (2009-2013)

Financiering: KU Leuven Onderzoeksfonds

Partners: KU Leuven (ESAT)

Beschrijving: Het doel van dit project is het verkennen van datagedreven, zwak gesuperviseerde en ongesuperviseerde methoden voor hiërarchische voorwerpsherkenning in spraak- en beeldherkenning. We onderzoeken wat direct uit data kan worden geleerd, op een bottom-up manier, door terugkerende patronen en structuren op meerdere abstractieniveaus te ontdekken. De belangrijkste uitdaging bestaat er dan in om uit te zoeken welke variaties relevant zijn en welke niet. We veronderstellen dat patronen die consequent gevonden worden over meerdere modaliteiten waarschijnlijk de semantisch meer betekenisvolle en interessante zijn. Daarom is de integratie van verschillende modaliteiten essentieel om te slagen in een dergelijke ongesuperviseerde of zwak gesuperviseerde aanpak. Door het koppelen van uitingen aan visuele waarneming en vice versa, komen we dichterbij semantisch begrijpen van beeld en spraak, zoals de mens dit leert.

Output:

- doctoraat Meng Sun
- databank van kleuterserie geannoteerd in de visuele en auditieve modaliteit

12. Titel: SIF: Subword Information Fusion in Exemplar Based Recognition (2011-2015)

Financiering: KU Leuven (ESAT)

Partners: KU Leuven (ESAT)

Beschrijving: In dit project wordt de state-of-the-art op het gebied van voorbeeldgebaseerde spraakherkenning verder ontwikkeld. In het verleden werd informatiefusie uitgevoerd op woordniveau, wat goed werkt voor korte woorden, maar naarmate het aantal lettergrepen toeneemt gaat dit minder goed. In dit project voegen we informatie samen op subwoordniveau en doen we de verwerking met subwoordeenheden in plaats van woorden (syllaben, demi-syllaben, morfemen, ...). Tegelijkertijd onderzoeken we ook extra K-NearestNeighbour gebaseerde kenmerken die kunnen worden toegevoegd aan het fusieproces.

13. Titel: AMiCA: Automatic Monitoring for Cyberspace Applications (2013-2016). Zie onder

14. Titel: Gezamenlijke Scheiding en Identificatie van Geluidsbronnen (2017-2021)

Financiering: FWO Strategische Onderzoeksbeurs

Partners: KU Leuven (ESAT)

Beschrijving: Wanneer mensen gelijktijdig spreken of wanneer een achtergrondspreeker extra ruis introduceert, wordt de verwerking van het spraaksignaal veel moeilijker. Ook is het bij meerdere sprekers vaak gewenst om hun identiteit te bepalen. De twee belangrijkste onderzoeksdoelstellingen van het project zijn bronscheiding (schatting van het origineel spraaksignaal van elke persoon) en sprekeridentificatie (wie spreekt er). Deze problemen kunnen het best gezamenlijk worden opgelost. Daardoor helpen ze elkaar constructief om een betere performantie te bekomen. Sequentiële benaderingen zijn minder efficiënt, omdat elke stap afzonderlijk wordt geoptimaliseerd waarbij de andere stap verwaarloosd wordt. Bovendien zijn vaak zowel bronscheiding als sprekerherkenning tegelijk noodzakelijk zoals bijvoorbeeld bij automatische transcriptie van vergaderingen.

Output: doctoraat van Jeroen Zegers (binnen enkele jaren)

15. Titel: Betekenis geven aan geluid (2017-2021)

Financiering: FWO aspirantenmandaat

Partners: KU Leuven (ESAT)

Beschrijving: Dit project breidt de huidige state-of-the-art uit op het gebied van audio-analyse, die momenteel voornamelijk gericht is op classificatie van specifieke, geïsoleerde geluiden die typisch zijn opgenomen in een kleine set omgevingen. Het doel is om cross-modale geluidsrepresentaties te leren die het geluid kunnen associëren met een realisatie in een andere modaliteit, in het bijzonder tekst of beeld. We gebruiken daarbij geen databank met handmatig gelabelde klankevents, maar we leren uit grote collecties zwak geannoteerde multimodale data waar we uit de context de relatie kunnen afleiden tussen de modaliteiten. De klankopnames bevatten geen geïsoleerde, maar gelijktijdige of (in de tijd) overlappende geluiden zoals ze in de realiteit voorkomen.

Output: doctoraat van Wim Boes (binnen enkele jaren)

16. Titel: What's in the Bag for Latent Variable Language Modeling? (2014-2018)

Financiering: KU Leuven - Radboud joint doctoral program.

Partners:

- a. KU Leuven (ESAT) (BE)
- b. Radboud Universiteit (CLST) (NL)

Beschrijving: Een van de grootste problemen bij de natuurlijke taalverwerking is cross-domain taalmodellering. Als we op medische teksten trainen en medische teksten evalueren, verwachten we een redelijke prestatie. Als we echter op medische teksten trainen en wetteksten evalueren, gaan de prestaties achteruit. Hoewel dit voor de hand ligt, zien we hetzelfde gedrag in meer gelijkaardige scenario's zoals Wikipediateksten en krantenteksten. We zoeken het antwoord in de ontwikkeling van rijke voorbeeldgebaseerde taalmodellen die op een statistische en impliciete manier, regelmaat ontdekken in onderliggende woordsequenties. We willen dit doen door rijke woordvoorstellingen te combineren met latent variabele modellen.

Output:

- tools voor Bayesiaanse taalmodellering op GitHub
- doctoraat Louis Onrust (nog af te leggen)

17. Titel: Ruisrobuuste automatische spraakherkenning gebaseerd op spectro-temporele kenmerken (2012-2018)

Financiering: eigen middelen

Partners:

- a. KU Leuven (ESAT) (BE)
- b. University of Szeged (HU)

Beschrijving: Spraakherkenning lijdt nog altijd performantieverlies in de aanwezigheid van storende ruis en het onderzoek naar robuuste spraakherkenning is nog altijd actueel. In spectro-temporele verwerking is het doel om tegelijkertijd de spectrale en temporele modulaties van het signaal vast te leggen. De resulterende kenmerken worden dan op dezelfde manier gebruikt als traditionele kenmerken zoals MFCCs. In dit project onderzoeken we verschillende methoden om spectro-temporele verwerking te integreren in een spraakherkenningssysteem en tonen we aan dat deze inderdaad leiden tot een robuustere spraakherkenning en resultaten leveren die competitief zijn met de state-of-the-art.

Output: doctoraat György Kovács (binnenkort af te leggen)

18. Titel: COGNIMUND (Cognitief begrijpen van beeld) (2010-2015)

Financiering: ERC Starting Independent Researcher Grant Tinne Tuytelaars

Partners: KU Leuven (ESAT-Visics)

Beschrijving: In dit project werden volgende problemen bestudeerd: 1) ontwerp van innovatieve en meer robuuste representaties van beeld die de intrinsieke structuur van beeld integreren, en 2) ontwikkeling van zwakgesuperviseerde methoden die leren uit een multimodale input (voornamelijk uit de combinatie van beeld en tekst).

Output: publicaties

19. Titel: NEFOLOGISCHE SEMANTIEK: Distributionele attestatiewolken als betekenisdiagnostiek in taalvariatieonderzoek (2015-2021)

Financiering: BOF (3H150305)

Partners: KU Leuven (QLVL)

Beschrijving: Met het toenemende belang van corpusmateriaal in de taalkunde, stijgt ook de nood aan aangepaste methodes om betekenisinformatie aan het corpusmateriaal te onttrekken. In het hier voorgestelde onderzoek wordt de distributionele corpussemantiek verfijnd in de vorm van een semantische diagnostiek gebaseerd op token clouds ('attestatie wolken'), d.w.z. in een meerdimensionele vectorruimte identificeerbare clusters van distributioneel vergelijkbare attestaties van woorden of uitdrukkingen. Deze 'wolkenkundige' methode voor het terugvinden van talige betekenissen in corpusmateriaal wordt in de eerste fase van het project verfijnd, en wordt in de tweede fase van het project ingezet in twee descriptieve onderzoeklijnen: lectometrisch onderzoek van de relatie tussen taalvariëteiten, en corpusgebaseerd variationistisch grammaticaonderzoek.

Output: publicaties

20. Titel: TermWise (Hulpbronnen voor gespecialiseerd taalgebruik) (2009-2013)

Financiering: KU Leuven, IOF-kennisplatform.

Partners:

- a. KU Leuven (QLVL) - coördinator
- b. KU Leuven (CW-LIIR)
- c. KU Leuven (De Nayer)
- d. KU Leuven (Lessius)

Beschrijving: Er is een groeiende behoefte aan effectieve meertalige oplossingen die zakelijke en professionele communicatie ondersteunen. Het kennisplatform TermWise is een multidisciplinaire samenwerking met de expliciete doelstelling om gebruikersgericht onderzoek naar meertalige taaltechnologieën te verrichten. Meer specifiek wil het platform software ontwikkelen die taalprofessionals, zoals vertalers of copywriters, helpt effectiever om te gaan met gespecialiseerde teksten. Deze teksten, bijvoorbeeld medische of juridische documenten, zijn vol domeinspecifiek jargon en terminologie. TermWise streeft naar automatische extractie van meertalige woordenboeken en vertaalgeheugens uit parallelle en vergelijkbare corpora die zijn geschreven in het Nederlands en het Frans.

Output:

- Dit project leidde tot de ontwikkeling van een geïntegreerd platform voor meertalige termextractie. Deel hiervan is o.a. een model voor termvertaling op basis van probabilistische grafische modellen, dat getraind werd op vergelijkbare corpora en toegepast op cross-talige retrieval. Er werden ook verschillende modellen gebouwd voor hoogkwalitatieve woordvertaling op basis van vergelijkbare corpora en parallelle corpora.
- Doctoraten Ivan Vulic en Dirk De Hertog

Outreach: Deze technologieën zijn geïntegreerd in een platform voor meertalige terminologie-extractie, ontwikkeld door alle consortiumpartners. Deze software is vrij beschikbaar.

21. Titel: Impact van het gebruik van corpora en CAT tools op de kwaliteit en efficiëntie van vertaalprocessen en vertalingen van gespecialiseerde teksten (2016-2018)

Financiering: STG-project (STG/15/040/IOF)

Partners: KU Leuven (QLVL)

Beschrijving: In een algemene Europese context werd recent onder professionele vertalers een beperkte kennis vastgesteld van de voordelen van het gebruik van (gespecialiseerde) corpora. Dit project stelt in een eerste fase tot doel het huidige gebruik van corpora door professionele vertalers in Vlaanderen te documenteren, naast de meer algemeen gebruikte tools voor computerondersteund vertalen (CAT). In een tweede fase zullen deze data de basis vormen voor een taalkundige analyse van de mogelijke effecten van het gebruik van corpora en CAT tools op de kwaliteit van vertaalde teksten, met inbegrip van criteria voor de algemene tekstuele gepastheid van vertaalde teksten in doeltaal en –domein. Het project focust op gespecialiseerde (wetenschappelijke, medische, technische en juridische) vertalingen (Engels/Nederlands).

Output: publicaties

22. Titel: Latijnse autoriteit en constructionele transparantie aan het werk: neologismen in het Franse medische vocabularium van de middeleeuwen en hun voortbestaan. (2014-2018)

Financiering: KU Leuven (BOF-OT)

Partners:

- a. KU Leuven (QLVL)
- b. UC Louvain
- c. ATILF (Nancy, DMF2015)

Beschrijving: Dit project onderzoekt waarom bepaalde Franse neologismen, die tijdens de middeleeuwen in het domein van de geneeskunde ontstonden, konden voortbestaan, terwijl andere na verloop van tijd verdwenen. Onze hypothese luidt dat morfologie, in het bijzonder constructionele transparantie, op een cruciale manier bijdroeg tot het behoud van die woordenschat. Middeleeuwse medische neologismen, geselecteerd op basis van een gedigitaliseerd corpus, worden onderworpen aan een grondige morfologische analyse volgens een set van criteria, die vervolgens statistisch worden verwerkt om de belangrijkste parameters die een rol spelen in het behoud van de woordenschat bloot te leggen. Een elektronische morfologische database van medische termen wordt bovendien beschikbaar gesteld aan de onderzoeksgemeenschap.

Output:

- online morfologische database Latijn – Middelfrans
- doctoraat Céline Szeceł

23. Titel: A corpus linguistic approach of Classical Hebrew: a pilot project (Postdoctoraal project Johan de Joode) (2017-)

Financiering: KU Leuven C1

Partners:

- a. KU Leuven (QLVL)
- b. KU Leuven (FTRW)

Beschrijving: Dit project bestudeert de teksten van de Hebreeuwse Bijbel en de Dode Zeerollen aan de hand van computationeel stylistische methoden. De aandacht gaat in eerste instantie uit naar het clusteren van de documenten met als doel om inzicht te krijgen in hun stijl, aan de ene kant, en hun ontstaansgeschiedenis, aan de andere kant. Binnen de studie van klassiek Hebreeuwse teksten is er onafgebroken discussie over auteurschap, tekstuele verwantschap, en herkomst. Dit project beoogt de deelname aan deze debatten aan de hand

van specifieke stylometrische en corpus-linguïstische analyses die rekening houden met het effect van genre, citaten, spelling, document lengte, fragmentatie, en inhoud. Dit project kwantificeert, in de mate van het mogelijke, stijl in het klassiek Hebreeuws.

Ouput:

- Python-scripts en R-scripts
- nieuwe datasets Bijbels Hebreeuws en Qumran Hebreeuws

24. Titel: MUSTER (MULTimodale verwerking van ruimtelijke en temporele expressies) (2016-2019)

Financiering: CHIST-ERA EU

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. ETH Zurich (CH)
- c. University Pierre and Marie Curie (FR)
- d. University of the Basque Country (ES)

Beschrijving: Het project ontwikkelt technieken voor het automatisch begrijpen van natuurlijke taal en in het bijzonder het automatisch interpreteren van ruimtelijke en temporele expressies in de taal door te trainen met beeld (statisch beeld en video) en hun tekstuele beschrijvingen. Dit is een fundamenteel onderzoeksproject dat focust op het leren van representaties van de betekenis van taal die automatisch worden aangeleerd in de context van beeldmateriaal.

Output: belangrijke publicaties in vooraanstaande conferenties en tijdschriften, bijvoorbeeld Proceedings AAAI, TACL, etc.

Outreach: De ontwikkelde technologieën hebben impact op mens-machinecommunicatie in de robotica.

25. Titel: MUSE (Machinaal begrijpen voor interactieve storytelling) (2012-2016)

Financiering: EU FP7-296703 (FET-open call)

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. Josef Stefan Institute (SI)
- c. Teesside University (UK)
- d. Leiden University (NL)
- e. Haute Autorité de Santé (FR)

Beschrijving: Het MUSE-project introduceert een nieuwe manier om tekstuele informatie te verkennen en te begrijpen door "tekst tot leven te brengen" door middel van 3D-interactieve verhalen. MUSE verwerkt natuurlijke taal, vertaalt deze naar een formele kennisrepresentatie die acties, actoren, plots en de omringende wereld weergeeft door middel van geavanceerde technieken van machinaal leren, en genereert uit de representatie een virtuele 3D-wereld. Het lange termijn-doel van MUSE is het automatisch vertalen tussen taal en de visuele modaliteit mogelijk te maken. LIIR was betrokken bij de coördinatie van het project en bij het onderzoek naar geavanceerde natuurlijke taalverwerking.

Output: Een platform voor de natuurlijke taalverwerking van Engelstalige teksten werd gerealiseerd dat modules voor de natuurlijke taalverwerking omvat inclusief semantische rolherkenning, coreferentenresolutie, en modules voor het herkennen van temporele en ruimtelijke relaties.

Outreach: De resultaten gaven aanleiding tot voordrachten op uitnodiging op events georganiseerd door de Europese commissie en door RE•WORK, een belangrijk platform voor communicatie van toponderzoek in de artificiële intelligentie.

26. Titel: TOSCA-MP (Taak-georiënteerd zoeken en annotatie van content voor de productie van media (2011-2014))

Financiering: EU FP7-287532

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. Joanneum Research Forschungsgesellschaft (AT)
- c. playence KG (AT)
- d. Deutsche Thomson (DE)
- e. Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung (DE)
- f. Institut für Rundfunktechnik GmbH (DE)
- g. Union Européenne de Radio-Télévision, Switzerland (CH)
- h. VRT (Vlaamse Radio- en Televisieomroeporganisatie) (BE)
- i. RAI-Radiotelevisione Italiana SpA (IT)
- j. Fondazione Bruno Kessler (IT)

Beschrijving: TOSCA-MP heeft algoritmes gebouwd voor cross-media-annotatie en ontsluiting van video. De taken van LIIR betroffen geavanceerde multimodale informatie-extractie (met de nadruk op het ophalen van informatie uit de teksten van videotranscripties) en het ontwerpen van geschikte ontsluitingsmodellen voor het doorzoeken van heterogene content.

Output: LIIR heeft modules ontwikkeld voor videosegmentatie, aligneren van namen met gezichten en objecten, en cross-modale coreferentenresolutie.

Outreach: Het project TOSCA-MP behaalde de eerste plaats wat betreft technische impact en de tweede plaats wat betreft economische impact in de MaxiCulture Coordination Action of DigiCult.

27. Titel: TERENCE (Een adaptief leersysteem om te redeneren over verhalen voor kinderen met leesmoelijkheden en hun begeleiders) (2010-2013)

Financiering: EU FP7- 257410

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. Università degli Studi dell'Aquila (IT)
- c. Libera Università di Bozen-Bolzano (IT)
- d. Fondazione Bruno Kessler FBK (IT)
- e. Università degli Studi di Padova (IT)
- f. Università degli Studi di Verona (IT)
- g. Moholy-Nagy University of Budapest (HU)
- h. University of Sussex (UK)
- i. L3S Research Centre Hannover (DE)
- j. Biomedicine Intelligent Systems (ES)
- k. Centre for Scientific Visualisation (SI)
- l. SIVCO Romania (RO)

Beschrijving: TERENCE combineert natuurlijke taalverwerkingstechnieken met inzichten uit de onderwijspsychologie om een adaptief leersysteem te ontwikkelen voor kinderen van 8-10 jaar die moeite hebben om tekst te begrijpen. De natuurlijke taalverwerkingsmodellen identificeren automatisch de belangrijke gebeurtenissen en de temporele en causale relaties daartussen om een semantische weergave van het verhaal op te bouwen. Slimme games worden vervolgens bovenop deze semantische representaties geconstrueerd om kinderen te stimuleren te redeneren over de gebeurtenissen in het verhaal, waardoor hun geletterdheid wordt verbeterd.

Output: Het onderzoek heeft geresulteerd in een demonstrator die de gebeurtenissen, hun actoren en coreferenten, en temporele relaties herkent in een verhaal.

Outreach: Een deel van de onderzoeksresultaten werden gecommmercialiseerd door de Università degli Studi dell'Aquila (IT).

28. Titel: MetaHaven (Management van verrijkte content) (2016-2018)

Financiering: Digital News Initiative Innovation Fund Google

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. KU Leuven (ESAT-Visics) (BE)
- c. Google (USA)
- d. Roularta Media Group (BE)
- e. Zeticon (BE)

Beschrijving: MetaHaven annoteert automatisch nieuwsteksten, afbeeldingen en video met IPTC-metadata (International Press Telecommunications Council) om een betere toegang tot de nieuwsmedia te bieden. De nadruk ligt op het bouwen van geschikte multimodale representaties die het categoriseren van nieuwsdata vergemakkelijken.

Output: We hebben een model gebouwd dat beeld verrijkt met tekstuele informatie en dat de classificatie van beeld in IPTC-klassen vergemakkelijkt inclusief beelden zonder bijhorende tekst. De classificatie wordt verbeterd door het aligneren van persoonsnamen in tekst en gezichten in beeld.

Outreach: De modellen worden in de systemen van Roularta Media Group en Zeticon geïntegreerd.

29. Titel: DISK (Ontdekking van kennis over Chinese medicinale planten in biomedische teksten) (2012-2016)

Financiering: KU Leuven –Tsinghua BIL/012/008

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. KU Leuven (Animal Physiology and Neurobiology)
- c. Tsinghua University (CN)

Beschrijving: Het project betreft kennisextractie (entiteiten en hun relaties) uit biomedische teksten met betrekking tot de traditionele Chinese geneeskunde. Deze biomedische teksten vormen een kennisbestand dat in de afgelopen millennia is opgebouwd waarvan de oudste schriftelijke bronnen 2000 jaar oud zijn en die nu voor een groot deel gedigitaliseerd zijn.

Output: We hebben modules van semi-gesuperviseerd machinaal leren gebouwd en de integratie van latente factoren in deze modellen bestudeerd. De resultaten werden gepubliceerd in hoog aangeschreven tijdschriften.

Outreach: Een deel van de resultaten werd gepresenteerd tijdens de internationale workshop “Knowledge Discovery of Big Text: Challenges and Opportunities when Mining Biomedical Tekst” gehouden op 18 mei 2015 in de Faculty Club van de KU Leuven.

30. Titel: DAME (Een platform voor benaderingen van deep learning voor multilinguale tekstverwerking (2017-2019))

Financiering: KU Leuven C32/017/016

Partners: KU Leuven (CW-LIIR)

Beschrijving: We implementeren een grotendeels taalafhankelijk platform dat deep learning-software voor natuurlijke taalverwerking integreert en onderzoeken de eenvoudige overdracht van de modellen naar andere talen waarvoor slechts een beperkt aantal handmatig geannoteerde trainingsvoorbeelden beschikbaar zijn.

Output: We hebben een taalafhankelijke semantische rolherkenner ontwikkeld. Andere modules (bijv. coreferentenresolutie en herkenning van temporele expressies en relaties) zullen worden ontwikkeld.

Outreach: Platform en modules worden na publicatie publiek beschikbaar gesteld.

31. Titel: ACCUMULATE: ACquiring CrUcial Medical information Using LAnguage TEchnology (2016-2019)

Financiering: IWT (SBO 150056)

Partners:

- a. KU Leuven (CW-LIIR) - coördinator
- b. KU Leuven (ESAT-STADIUS)
- c. KU Leuven (UZ Leuven)
- d. UAntwerpen (CLiPS)
- e. UAntwerpen (UZA Antwerpen)

Beschrijving: ACCUMULATE extraheert kennis uit de vrije tekst van klinische rapporten geschreven in het Engels en het Nederlands met behulp van technieken van machinaal leren. Hiertoe zal gewerkt worden aan het ontwerp, de ontwikkeling en de evaluatie van geavanceerde taaltechnologiecomponenten voor diepe semantische analyse van teksten die vaak morfosyntactisch niet welgevormd zijn. Een bijkomende focus ligt op de gemakkelijke herbruikbaarheid van de ontwikkelde technologie in andere domeinen en talen en op het gebruik van visualisatietechnieken.

Output: We hebben nieuwe leermodellen ontwikkeld voor het herkennen van medische events en temporele relaties tussen deze events. We onderzoeken momenteel spellingcorrectie en tekstnormalisatie voor klinische tekst

- software voor event detection in klinische tekst
- software voor analyse van negatie en modaliteit in klinische tekst.
- nieuwe leermodellen voor het herkennen van medische events en temporele relaties tussen deze events
- momenteel wordt de extractie van ruimtelijke relaties tussen symptomen en lichaamsdelen bestudeerd in multimodale (tekst en beeld) klinische radiologieverslagen. Daarnaast worden ook algoritmen bestudeerd voor hiërarchische classificatie van klinische rapporten met ICD-codering (International Classification of Diseases).

Outreach: De resultaten worden bekendgemaakt in de internationale literatuur en via een uitgebreide industriële adviesraad.

32. Titel: MARS (Machinaal lezen van patiëntendossiers) (2015-2018)

Financiering: KU Leuven C22/015/01

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. KU Leuven (QLVL) (BE)
- c. KU Leuven (UZ Leuven) (BE)
- d. KU Leuven (CiTiP) (BE)

Beschrijving: MARS extraheert informatie uit de vrije tekst van patiëntendossiers geschreven in het Engels en het Nederlands. De focus ligt op de terminologie- en informatie-extractie uit teksten die morfo-syntactisch niet welgevormd zijn.

Output: De output is gelijkaardig aan het ACCUMULATE-project.

Outreach: Het doel is een vervolproject met de industrie te realiseren.

33. Titel: Sunshine (Virtuele data-analist voor interactieve data-gedreven journalistiek) (2014)

Financiering: iMinds-MIX

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. KU Leuven (CW-HCI-Ariadne) (BE)

- c. Made in Wonder (BE)
- d. Bits of Love (BE)
- e. Mediahuis (BE)

Beschrijving: Sunshine heeft een interface en taaltechnologiemodule gebouwd om Nederlandstalige queries te vertalen naar een SQL-query die een relationele databank ondervraagt.

Output: Jan De Belder (onderzoeker van SunShine) heeft een iStart-beurs verkregen om de technologie verder te ontwikkelen in een start-up.

Outreach: Het team heeft een succesvolle demonstratie gegeven van de ontwikkelde technologie tijdens de iMinds-conferentie op 23 oktober 2014.

34. Titel: PARIS: Personalised Advertisements built from web Sources (Gepersonaliseerde advertenties gebouwd uit webbronnen) (2012-2016)

Financiering: IWT (SBO 110067)

Partners:

- a. KU Leuven (HCI-LIIR) -coördinator
- b. KU Leuven (HCI-Ariadne)
- c. KU Leuven (Visics)
- d. UGent (CWI)
- e. UGent (LT3)
- f. UAntwerpen (IELSP)
- g. Vlerick

Beschrijving: Sociale netwerken bevatten een schat aan informatie. PARIS onderzoekt technieken van de natuurlijke taal- en de beeldverwerking en van netwerkanalyse voor dit medium. Dit is een uitdaging omdat deze zogenaamde user generated content is geschreven in een taal die morfosyntactisch niet goed gevormd of bewaard is in beelden die niet professioneel zijn opgenomen. Een belangrijke toepassing betreft het bouwen van gebruikersprofielen op basis van deze informatie en het genereren van gepersonaliseerde reclame die past in de leefwereld van de gebruikers van de sociale media. Deze personalisatie moet in overeenstemming blijven met de wettelijke vereisten met betrekking tot de bescherming van persoonsgegevens en intellectuele eigendomsrechten.

Output: De output betreft de redactie van een boek getiteld Mining User Generated Content in 2014 uitgegeven door CRC Press (Taylor & Francis Group). We hebben interessante resultaten verkregen met betrekking tot de herkenning van leeftijd, geslacht en persoonlijkheid van gebruikers van sociale media op basis van probabilistische grafische

modellen en modellen van deep learning. Een ander resultaat betreft het cross- en multimodaal zoeken van modeartikelen in e-commerce.

Outreach:

- een publieke demo rond cross-modaal zoeken van modeartikelen (http://roshi.cs.kuleuven.be/multimodal_search/#/demopage)
- een publieke demo i.v.m. sentimentanalyse
- presentatie van de resultaten op de 44e jaarlijkse conferentie van EMAC - European Marketing Academy in Leuven, België van 26 mei tot 29 mei 2015 en tijdens een afsluitende workshop op 10 mei 2016
- de sentimentanalysesoftware ontwikkeld in het PARIS-project werd verder gecommercialiseerd in een VLAIO-innovatieproject voor het bedrijf Hello Customer

35. Titel: SAEL-RMR (Een synergistische benadering voor extractie, leren en redeneren bij het machinaal lezen) (2012-2017)

Financiering: FWO G.0356.12

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. KU Leuven (DTAI) (BE)
- c. UGent (Toegepaste wiskunde) (BE)

Beschrijving: Het doel van dit fundamenteel onderzoeksproject SAEL-RMR is om een synergie te vinden tussen het extraheren van informatie uit tekst, redeneren met deze informatie en machinaal leren.

Output: Het onderzoek was vooral gericht op entiteit-relatie herkenning in biomedische teksten met behulp van semi-gesuperviseerd machinaal leren, relationeel machinaal leren en de integratie van latente factoren in de modellen van machinaal leren.

Outreach: De outreach betreft publicaties onder andere in BioNLP, BIOSTEC en BMC Informatics, en het verwerven van een tweede plaats in de internationale competitie BioNLP 2013.

36. Titel: Radical (Radicale innovatie, competitie, coöperatie en groei) (2011-2016)

Financiering: KU Leuven GOA 12/003

Partners:

- a. KU Leuven (CW-LIIR) (BE)
- b. KU Leuven (ECOOM) (BE)

Beschrijving: Het hoofddoel van het project "Radical Innovations" was meer inzicht te krijgen in hoe "radicale innovaties" hun oorsprong vinden. Geautomatiseerde inhoudsherkenning in tekst (in octrooien, wetenschappelijke publicaties en zakelijk nieuws) levert hiervoor een belangrijke reeks indicatoren.

Output: We hebben software ontwikkeld voor de hiërarchische classificatie van patenten, wat een uitdagende taak is gezien de grote hoeveelheden classificatiecodes (in de duizenden) en het beperkte aantal met classificatiecodes geannoteerde trainingsvoorbeelden.

Outreach: Het project resulteerde in veel geciteerde publicaties.

37. Titel: Complement Raising and Cluster Formation in Dutch, A Treebank-supported Investigation (2011-2015)

Financiering: FWO

Partners: KU Leuven (CCL)

Beschrijving: De beschikbaarheid van treebanks voor gesproken en geschreven Nederlands heeft nieuwe en interessante mogelijkheden gecreëerd voor het empirisch onderzoek van de Nederlandse syntaxis. Dit project wil hun potentieel onderzoeken voor de studie van een typisch Nederlands fenomeen: de vorming van werkwoordclusters, d.w.z. constructies waarin meerdere werkwoorden samenkomen.

De studie heeft niet alleen geleid tot een beter begrip en een empirisch gefundeerde beschrijving van dit fenomeen, maar werpt ook licht op de meer algemene methodologische kwestie van hoe de datagedreven benadering van de treebank-mijnbouw en de op kennis gebaseerde benadering van de theoretische taalkunde het meest winstgevend kan zijn gecombineerd.

Output:

- doctoraat Liesbeth Augustinus
- verbeteringen aan GrE TEL

Outreach: taalkundigen (ook al door gebruik van GrE TEL)

38. Titel: Theory meets Quantity: Corpus-based Comparative Syntax (Theorie en kwantiteit: corpus-gebaseerde comparatieve syntaxis) (2016-2019)

Financiering: FWO

Partners: KU Leuven (CCL)

Beschrijving: Dit project betreft onderzoek naar het gebruik van treebanks, syntactisch geannoteerde corpora, voor de vergelijkende studie van verschillende talen.

Er is een snel groeiende gemeenschap van linguïsten die treebanks gebruiken voor hun onderzoek. Wat echter nog in de kinderschoenen staat, is het gebruik van treebanks voor de vergelijkende studie van verschillende talen. De corpusstudie die in dit project wordt uitgevoerd, zal niet worden beperkt tot het gebruik van eentalige corpora en treebanks. Een belangrijke innovatie omvat de exploitatie van de recent gebouwde Europarl parallelle treebank voor vergelijkende syntaxis. In een aantal casestudies zal worden aangetoond hoe de nieuwe hulpmiddelen en methoden kunnen leiden tot nieuwe inzichten en meer accurate verschijnselen in de West-Germaanse syntaxis, d.w.z. de vorming van werkwoordclusters in het Duits, Nederlands en Afrikaans.

Output: verbeteringen aan GrETEL 2.0, GrETEL 4 Afrikaans, PolyGrETEL

Outreach: taalkundigen

39. Titel: PICTo: Automatic translation between pictogram languages and Dutch (Een systeem voor automatische vertaling tussen pictogramtalen onderling en tussen pictogramtalen en het Nederlands) (2012-2014)

Financiering: Fonds M.M.Delacroix

Partners:

- a. KU Leuven (CCL)
- b. Thomas More Kempen
- c. WAI-NOT

Beschrijving: Het bevorderen van communicatiemogelijkheden / sociale inclusie voor mensen met leesproblemen, vooral gebruikers van de pictogrammen beschikbaar in WAI-NOT. Bij de start van het project werd er geen taaltechnologie gebruikt, enkel wanneer de naam van een pictogram toevallig identiek was aan het opgegeven woord, kon er een pictogram worden getoond. Maar zeker met homoniemen kon geen rekening worden gehouden.

Binnen het project is er een vertaalsysteem (Taal2Pictogram) tussen het Nederlands aan de ene kant en Beta/Sclera aan de andere kant. Door het introduceren van taaltechnologie (tokenization, basis spellingcorrectie, PoS-tagging, invoegen van zinsgrenzen, herkennen van

scheidbare werkwoorden, etc.), plus het koppelen met Cornetto/Open Source Dutch WordNet is er een systeem ontwikkeld dat ook rekening houdt met context en, in bepaalde gevallen, met de herkenbaarheid van een pictogram.

Output:

- Text2Picto (basisversie)
- koppeling Beta en Sclera aan de Cornetto en Open Source Dutch WordNet databanken
- aanpassingen voor Vlaams (genoemde databanken gaan uit van het Nederlands zoals in Nederland gebruikt)
- Sclera2Cornetto en Beta2Cornetto

Outreach:

- mensen met een verstandelijke beperking en/of behoefte aan ondersteunde communicatie
- WAI-NOT
- tools o.a. gebruikt in WAI-NOT (aangepaste, veilige website voor mensen met verstandelijke beperking)

40. Titel: Able to Include (2014-2017)

Financiering: EU (CIP PSP)

Partners:

- a. KU Leuven (CCL), België
- b. The National Microelectronics Applications Centre (MAC), Ierland
- c. Thomas More Kempen, België
- d. Universitat Pompeu Fabra, Spanje
- e. Building Bridges Training, UK
- f. Inclusion Europe, België
- g. Fundación Prodis, Spanje

Beschrijving: Dit project probeert de leefomstandigheden van mensen met een intellectuele beperking te verbeteren, onder meer door hen ook toegang te geven tot de social media (e-inclusion). Dit door het bijvoorbeeld mogelijk te maken met pictogrammen (Sclera en Beta) te communiceren, al dan niet in combinatie met een natuurlijke taal, in dit geval Nederlands, Engels of Spaans. De opzet maakt het mogelijk andere talen toe te voegen, mits er een WordNet voor bestaat, en ook andere sets pictogrammen, zowel die met in principe één concept per pictogram als die met meerdere.

Partner CCL maakte in dit project de Picto-technologie taalonafhankelijk en breidde de taaltechnologische ondersteuning uit om ook Engelse en Spaanse input te verwerken (taggers, lemmatizers, maar ook voor taalspecifieke fenomenen, zoals prodrop (Spaans)). Alle gebruikte wordnets werden gekoppeld aan Princeton WordNet. Een aantal eindgebruikers gaven steeds suggesties en feedback.

Output:

- AbleSocial, AbleChat, AbleMail (Kolumba): deze apps maken essentieel gebruik van Picto2Text en Text2Picto 2.0 ontwikkeld in "Pictographic Communication Technologies for Browsing the Web", ook een project van KU Leuven (CCL)
- WAI-NOT maakt nu onder meer gebruik van de nieuwe versie van Text2Picto, en van Picto2Text.

Outreach:

- De applicatie wordt gebruikt door mensen met nood aan ondersteunde communicatie, zoals mensen met een intellectuele beperking, migranten, prelinguaal doven, ouderen, laaggeschoolde moedertaalgebruikers,...
- Publiciteit, buiten de kring van computationele linguïsten:
 - o Language Industry Award voor "Best Language Service 2015"
 - o Best Paper Award - Web for All Conference (April 2017, Perth, Australia) - met Universitat Pompeu Fabra
 - o Pioniersprijs Humane Wetenschappen 2017 - Honorary Mention (May 2017, Leuven)
 - o Mediawijs: M-Award (September 2017, Brussels) - met Thomas More, Kempen
 - o The Big Draw: voor mensen met een cognitieve beperking een les aan de KU Leuven over vertaaltools tussen pictogrammen <=> Nederlands
 - o Artikel in Able to Include Newsletter
 - o Artikel op Facebookpagina van Inclusion Europe
 - o Een Pictogram Zegt Meer dan Duizend Woorden - artikel voor Triptiek (tijdschrift van Downsyndroom Vlaanderen)
 - o Artikel in Campuskrant KU Leuven (ook online)
 - o Computational Linguist Leen Sevens Talks to Durtti (gelinkt aan Artificial Intelligence groep op LinkedIn)
 - o Text2Picto sloopt muren. Artikel in De Taalsector over Text2Picto en de Falling Walls Lab
- Organisatie ISI-NLP workshop (LREC 2016)
- Organisatie ISI-NLP2 workshop (LREC 2018)

41. Titel: Pictographic Communication Technologies for Browsing the Web (Pictogramvertaaltechnologie voor Webnavigatie) (2015-2018)

Financiering: IWT

Partners: KU Leuven (CCL)

Beschrijving: In dit project werden Picto2Text en Text2Picto verder uitontwikkeld, met vooral aandacht voor het Nederlands, en minder voor andere talen. Omdat pictogrammen onder meer geen functiewoorden en morfologische kenmerken bevatten, of een pictogram zowel

als werkwoord of als zelfstandig naamwoord kan worden gebruikt, is het genereren van acceptable zinnen niet eenvoudig. Een ander issue is het feit dat een pictogram kan corresponderen met meerdere woorden in de zin.

Er werd enerzijds gewerkt met verschillende taalmodellen (gebruikmakend van n-grammen of van Long Short-Term Memory (neurale netwerken)), anderzijds met MT-technieken zoals Phrase-based Machine Translation (Moses) en Neural Machine Translation (OpenNMT). Ook in dit project staat de (eind)gebruiker centraal (user-centered approach).

Output:

- doctoraat Leen Sevens (verwacht in december 2018)
- Text2Picto 3.0, Picto2Text 2.0, spellingcorrector, simplifier, semantische desambiguator

Outreach:

- Falling Walls Lab - Leuven 2017 (Leen Sevens: Eerste prijs van de jury + prijs van het publiek. Mag deelnemen aan de finale in Berlijn (100 kandidaten uit de hele wereld)
- Falling Walls Lab - Berlijn 2017
- filmpje op YouTube (deelname Berlijn)
- 'Lanakense vindt systeem uit ...', Belang van Limburg, n.a.v. deelname Berlijn
- "Picto", Interview Radio 1, Nieuwe Feiten. Bekende VRT'ers testen het systeem uit
- "Lanakense ontwikkelt pictogrammentaal ...", interview TV Limburg
- interview: Leen Sevens Onderzoek in de kijker' (KU Leuven)
- "Kom naar buiten met je onderzoek", artikel Leen Sevens voor "KU Leuven Blogt"
- Leen Sevens was ook medewinnaar van de awards genoemd onder 'Able to Include'

42. Titel: CLARIN-PLUS (2016)

Financiering: EU

Partners:

- a. KU Leuven (CCL)
- b. CLARIN-ERIC

Beschrijving: KU Leuven (CCL) heeft in het kader van CLARIN-Plus een workshop georganiseerd om bepaalde groepen onderzoekers, binnen en buiten de CLARIN-community, te tonen wat de in-line mogelijkheden zijn met de beschikbare tools en resources. In Leuven betrof het onderzoekers met interesse voor digitale collecties krantenmateriaal.

Output: driedaagse workshop "Digital Collections of Newspapers" (video beschikbaar)

Outreach: onderzoekers humane en sociale wetenschappen

43. Titel: Nederbooms: Exploitation of Dutch treebanks for linguistic research (Exploitatie van Nederlandse treebanks voor taalkundig onderzoek) (2010-2012)

Financiering: EWI

Partners: KU Leuven (CCL)

Beschrijving: Het doel van dit project was gebruiksvriendelijke en goed gedocumenteerde hulpmiddelen te ontwikkelen voor de exploitatie van treebanks door taalkundigen die niet bekend zijn met taaltechnologie. Dit resulteerde in de eerste versie van GrETEL, een tool die treebank-bevraging mogelijk maakt door middel van een natuurlijk taalvoorbeeld in plaats van een formele zoekinstructie.

Output: GrETEL

44. Titel: Afribooms: Building a treebank for Afrikaans (2013-2014)

Financiering: NTU (voor het Vlaamse gedeelte)

Partners:

- a. KU Leuven (CCL)
- b. Noordwest-Universiteit / Noordwes-Universiteit, Zuid-Afrika

Beschrijving: Hoewel er een breed scala aan bronnen en technologieën bestaat voor het Nederlands, waardoor geavanceerde taaltechnologieën en geavanceerd taalkundig onderzoek kunnen worden geïmplementeerd, wordt het Afrikaans nog steeds beschouwd als een schaarse taal. Het doel van dit project was de oprichting van een Afrikaanse treebank. De AfriBooms treebank is online via GrETEL opzoekbaar voor het Afrikaans. CCL heeft vooral gewerkt aan de aanpassingen die nodig waren om de tools die voor het Nederlands beschikbaar waren, om te zetten naar tools voor het Afrikaans.

Output: GrETEL 4 Afrikaans

45. Titel: GrETEL2.0: Querying very large treebanks (2013-2014)

Financiering: NTU (behalve voor Nederlandse deel educatieve module)

Partners: KU Leuven (CCL)

Beschrijving: Het doel van dit project was de aanpassing van GrETEL voor het bevragen van zeer grote treebanks, zoals het SoNaR-referentiecorpus voor geschreven Nederlands (ca. 500M tokens, 41M-zinnen). De grootste uitdaging was om deze enorme treebank net zo snel doorzoekbaar te maken als de kleine treebanks die al in GrETEL waren opgenomen. Er is ook een educatieve module MIMORE-GrETEL gemaakt, samen met het Meertens Instituut (prof Sjef Barbiers) om onderzoekers en studenten te laten zien hoe twee zeer verschillende tools, nl. MIMORE (dialecten) en GrETEL (standaardtaal) in combinatie kunnen worden gebruikt om hetzelfde fenomeen in beide typen taal te onderzoeken.

Output:

- GrETEL 2.0
- In Utrecht (Jan Odijk en collega's) is er veel interesse voor GrETEL. Er is daar verder gewerkt op GrETEL 2 en 3, resulterend in GrETEL 4. Het nu bv. ook mogelijk eigen data en metadata op te laden, en gericht de zoekresultaten te doorzoeken, onder meer gebruikmakend van metadata.

Outreach:

- Steven Krauwer Award for young Scientists (CLARIN)
- taalkundigen, studenten

46. Titel: ReGround: relationeel aarden van symbolen via het leren van affordanties (2015-2020)

Financiering: FWO (CHIST-ERA)

Partners:

- a. KU Leuven (DTAI)
- b. Orebro University (Zweden)
- c. Koc University (Turkije)

Beschrijving: Het symbool-aardingprobleem probeert om symbolen uit taal te associëren met een corresponderend element in de omgeving. Traditioneel heeft research zich gericht op het identificeren van individuele objecten en hun eigenschappen. Op een gelijkaardige manier ligt de focus binnen de robotica voornamelijk op het leren van affordanties voor individuele objecten. Deze benaderingen nemen vaak de volledige context van de omgeving niet in overweging. Daarnaast is het ook zo dat de toestand waarin de omgeving zich bevindt bepaalde acties toelaat. Dit project stelt de hypothese dat het aardingsproces de volledige context van de omgeving in overweging moet nemen om in staat te zijn symbool-aarding uit te voeren, en om beter te reageren op een nieuwe taal en omgeving. Dit project is gericht op het ontwikkelen van een nieuwe aanpak voor aarding die het optilt tot het relationele niveau, waar een actor kan redeneren over de relaties tussen meerdere symbolen in de taal, en tussen meerdere elementen in de omgeving, en waar het in staat is om affordanties te leren die de relatie tussen objecten en hun eigenschappen, acties en omgeving bevatten. Dergelijke relaties leren vereist het combineren van informatie uit verschillende modaliteiten zoals taal en perceptie.

47. Titel: STARLING: Statistisch relationeel leren van natuurlijke taal (2010-2013)

Financiering: FWO

Partners:

- a. KU Leuven (DTAI)
- b. UAntwerpen (CLiPS)

Beschrijving: Dit project wil onderzoeken hoe technieken van statistisch relationeel leren gebruikt kunnen worden voor het leren van natuurlijke taal. Daarbij zal de nadruk liggen op uitdagende taken van de natuurlijke taalverwerking, zoals “semantic role labeling”, waarbij syntactische en semantische afhankelijkheden, gestructureerde en ongestructureerde gegevens, lokale en globale modellen, en probabilistische en logische informatie met elkaar moeten gecombineerd worden. Voor wat betreft het statistisch relationeel leren zal onze aandacht vooral uitgaan naar representaties die gebaseerd zijn op probabilistische uitbreidingen van de programmeertaal Prolog. Het project wil hierbij niet alleen betere resultaten in de natuurlijke taalverwerking realiseren maar ook betere algoritmen en systemen voor statistisch relationeel leren.

48. Titel: Trismegistos Project (2005-onbepaald)

Financiering: FWO, Hercules, BOF e.a. Zie http://www.trismegistos.org/about_funding.php

Partners:

- a. KU Leuven (Onderzoeksgroep Oude Geschiedenis)
- b. Universität zu Köln
- c. Databanken, zie ook http://www.trismegistos.org/about_partners.php

Beschrijving: Sinds 2005 coördineert de onderzoeksgroep Oude Geschiedenis onder leiding van professor Mark Depauw het digitale platform Trismegistos (TM). Deze online database bevat metadata over teksten uit de oudheid die dateren van ongeveer 800 v.Chr. tot 800 n.Chr. Oorspronkelijk inventariseerde TM alle personen met een titel uit Ptolemaïsch Egypte in een database: **TM People**. Deze personen worden geïdentificeerd aan de hand van een uniek TM-nummer, eenzelfde procedé dat ook wordt toegepast om teksten te onderscheiden (**TM Texts**). De database werd in de daaropvolgende jaren uitgebreid met een historische topografie van het Fajoemgebied in de Grieks-Romeinse periode van 300 v.Chr. tot 700 n.Chr. in **TM Places**, de identificatie van meer dan 300 papyrusarchieven uit Grieks-Romeins Egypte in TM Archives en een overzicht van alle papyruscollecties ter wereld in TM Collections. Hiernaast werden in de Leuven Database of Ancient Books de metadata verzameld voor alle restanten van Griekse, Latijnse en Koptische boeken tot 800 n. Chr.

Via samenwerkingen met verschillende partners slaagde Trismegistos erin om steeds verder uit te breiden en om ook informatie over teksten uit kleinere taalgroepen en gebieden buiten Egypte te integreren in TM.

Output: het platform Trismegistos, <http://www.trismegistos.org>

49. Titel: Harnessing collective intelligence in e-learning environments (2007-2012)

Financiering: KU Leuven – IOF

Partners: KU Leuven (ITEC)

Beschrijving: Bij Web 2.0-toepassingen is het mogelijk om in functie van het verbeteren van de bestaande webservice of het ontwikkelen van nieuwe services gebruik te maken van de 'collective intelligence', die via tracking en logging wordt opgebouwd door de gebruikers van de webservice. Een van de uitdagingen is de vraag hoe elektronische omgevingen zich automatisch kunnen aanpassen aan het geregistreerde profiel van de gebruiker. Dit project behandelt deze uitdaging vanuit drie perspectieven, telkens door één doctoraatsstudent, ondersteund door een informaticus en een statisticus: - het informatica-perspectief: het ontwikkelen van algoritmes voor het vertalen van de opgebouwde intelligentie naar instructies voor verschillende 'dragers' en 'contexten' - het data-analytisch perspectief: de automatische selectie van verschillende types taken, aangepast aan de vaardigheid van de leerder en aan de context- het onderwijstechnologisch perspectief: nagaan of en hoe adaptiviteit het leerrendement kan verhogen.

Output: doctoraten van Ahmet Soylu, Kelly Wauters, Mieke Vandewaetere

50. Titel: COBALT - Communicating and building bridges thanks to the Acquisition of Languages through Technologies (2008-2013)

Financiering: EU – Interreg

Partners:

- a. KU Leuven (ITEC) (coördinator)
- b. Université Lille 3
- c. Université de Mons
- d. WTV Zuid
- e. VDAB
- f. Forem

Beschrijving: COBALT beoogt de verdere ontwikkeling en vernieuwing van de elektronische taalleeromgeving Fanel. Fanel is gratis toegankelijk via het internet en biedt eenieder die zijn Frans of Nederlands wil bijspijkeren, interactieve leeractiviteiten aan, vertrekkend van levendig en authentiek beeldmateriaal van de regionale televisiezenders. Zo draagt Fanel bij tot de bevordering van de taalvaardigheid, de interculturele communicatie en de innovatie van het taalonderwijs in een regio waar de kennis van de taal van de buur cruciaal is voor de socio-economische ontwikkeling. COBALT laat Fanel verder groeien, en dit met een nieuwe kijk op de leeromgeving voor de leerder, voor de leerkracht (Studio Fanel) en de Mediatic databank met authentiek videomateriaal.

Output: <http://www.fanel.eu> ; <http://www.kuleuven-kulak.be/cobalt>

Outreach: COBALT werd verkozen tot strategisch project door het Interreg IV programma: <https://www.youtube.com/watch?v=GSAzgwrPyYs>

51. Titel: iRead+ - the intelligent reading companion (2012-2013)

Financiering: iMinds – ICO

Partners:

- a. KU Leuven (ITEC) (coördinator)
- b. KU Leuven (CUO)
- c. VUB (SMIT)
- d. UGent (MMLab)
- e. Roularta Media Group NV
- f. D!nk
- g. BLCC
- h. Sensotec
- i. Regionale Media Maatschappij
- j. Playlane
- k. VZW De Kade

Beschrijving: Het doel van het iRead+ project was om te onderzoeken hoe elektronische leesteksten verrijkt kunnen worden door informatie toe te voegen die het leesproces kunnen ondersteunen (bv. foto's, extra informatie over personen die vermeld worden in artikels, uitleg bij moeilijke woorden, etc.). Het doelpubliek van iRead+ zijn zowel algemene lezers als taalleerders en kinderen met dyslexie. De technologie die voor het project ontwikkeld werd laat toe om de teksten zo goed als automatisch te verrijken en dat voor alle doelpublieken die in de verschillende use cases aan bod komen. Zie ook <https://www.ibbt.be/en/projects/iread>.

Output:

- Drie technologiedemonstrators voor 3 verschillende use cases:
 - o Demonstrator voor verrijking van leeservaring op tablet

- o Demonstrator voor taalleerders: verrijking van leesteksten in vreemde taal en aanbieder van oefeningen die automatisch gegenereerd worden
- o Demonstrator voor leerlingen met leesproblemen (dyslexie)
- De ontwikkeling van een “enrichment pipeline” die het mogelijk maakt om tekstbestanden te verrijken met content- en woordspecifieke informatie zoals foto’s, informatie uit Wikipedia, woordenboeken, etc.
- De ontwikkeling van een automatische oefengenerator. De oefengenerator kan op basis van om het even welke tekst een reeks morfologische oefeningen ontwikkelen voor taalleerders. Het systeem genereert ook automatisch eenvoudige correctieve feedback.
- Publicaties en presentaties op internationale congressen

Outreach:

- Organisatie van slotevent voor betrokken partners (maart 2014)
- Verspreiding van leaflet rond impact van iRead+

52. Titel: Games @ School (2012-2013)

Financiering: iMinds – ICON

Partners:

- a. KU Leuven (CUO) coördinator
- b. KU Leuven (ITEC)
- c. KU Leuven (ICRI)
- d. UGent (MICT)
- e. VUB (SMIT)
- f. Larian Studios
- g. die Keure
- h. Belgacom
- i. Grin

Beschrijving: The G@S project aims to research and develop an open platform for the creation and distribution of educational games. The education sector sees gaming, because of its systematic and self-motivating character, as a key future learning aid. The G@S project researches the establishment of a platform for the easy creation and free distribution of user-generated educational game content. To achieve this, the project aims to innovate in four main areas: authoring, adaptive content and feedback, understanding of the user and distribution.

Output: <http://youtu.be/RXSIm29rMzk>

Outreach: slotevent georganiseerd op het Media & Learning congres (11 december 2013)

53. Titel: Games online for Basic Language Learning (2012-2014)

Financiering: EU – Lifelong Learning Programme

Partners:

- a. KU Leuven (ITEC)
- b. Radboud Universiteit Nijmegen (CLST) –coördinator
- c. Newcastle University
- d. Televic Education
- e. Council for Scientific and Industrial Research (CSIR)
- f. Meraka Institute (Zuid-Afrika)

Beschrijving: This project aims to provide youths and adults wishing to improve their basic language skills with access to materials for the development of communicative proficiency in French, Dutch and English through web-based minigames that support spoken interaction. Worldwide, and especially in Europe, millions of people have to learn foreign languages every year, or a second language which they need to integrate in socio-economic contexts. Still, some groups participate to a lesser extent, or lose interest in formal language education altogether. Moreover, in foreign language curricula, it may be argued that speaking practice and grammatical accuracy receive little attention nowadays. Educational mini-games are small, self-contained games which focus on well-defined learning topics, which are highly motivating, reusable, and cost-effective. Mini-games are particularly fit for the development of language skills at the lower end of the proficiency scale. Moreover, there is evidence that disadvantaged and low-skilled language learners profit most from such mini-games. Finally, mini-games can be easily and cheaply distributed via the web, which has become a primary source of entertainment, especially for the mentioned groups. This project aims to develop such mini-games in a user-centred way for the teaching of grammar in use, vocabulary and basic communicative skills in French, English and Dutch as a foreign or second language. Learning materials target the A2 level of the CEFR. Dedicated speech recognition technology will be put to use for stimulating speaking practice in Dutch and English. Target users include language learners, teachers, and institutions in adult and secondary education, as well as independent language learners at home. Through the motivational aspect of gaming, we especially hope to make an impact on the language learning experience of low-skilled learners.

Output: <http://www.gobl-project.eu/>

Outreach: presentaties voor leerkrachten in 2013 op Edushock leerfestival & Dag van het Engels & Duits (Eekhoutcentrum)

54. Titel: Videostreaming as an interactive didactic instrument in computer-assisted language learning (Video.de) (2012-2014)

Financiering: KUL – Onderwijsontwikkelingsfonds

Partners:

- a. KU Leuven (ITEC)
- b. KU Leuven (CHIL) – coordinator
- c. Thomas More Antwerpen (MIDI)

Beschrijving: Ontwikkeling van een online platform waarmee streaming videodata op verschillende niveaus ontsloten kunnen worden. Het centrale aspect van deze omgeving is de studentgeoriënteerde tool voor de annotatie van video's, waarmee studenten actief aan de slag gaan met het audiovisuele materiaal in een tweede taal. De studenten annoteren video's volgens een taakspecifiek schema, dat op voorhand wordt bepaald door de docent en dat kan bestaan uit parameters zoals woordenschat, grammatica en discursieve kenmerken. Het ontwerp van zowel de annotatietool als van het flexibele annotatieschema maakt zijn toepassing in verschillende academische taalopleidingen mogelijk. Het didactische potentieel van de tool wordt getest in een pilootstudie voor de academische opleidingen Duits binnen de KU Leuven (Antwerpen, Brussel, Leuven, Kortrijk).

Outreach: onderwijsseminarie (https://www.arts.kuleuven.be/fac_kalender/agenda119)

55. Titel: Unified Thesaurus - Feasibility study for a common, unified thesaurus for the Flemish media, archives and heritage sector (2013)

Financiering: iMinds MiX

Partners:

- a. KU Leuven (ITEC)
- b. UGent (MMLab)
- c. UGent (IBCN)
- d. Faro
- e. Corelio
- f. Concentra Media
- g. Media Innovatie Centrum
- h. De Persgroep
- i. ARGOS vzw Centre for art and media
- j. Packed vzw - Centre of expertise in digital heritage
- k. Amsab
- l. Roularta Media Group NV

- m. Sanoma
- n. ADVN
- o. De Vijver Media
- p. WTV & FocusTV
- q. Vlaamse Nieuwsmedia
- r. Move
- s. VRT

Beschrijving: Bij annotatie vormen vaste, gestructureerde labels uit thesauri de ankers. Een gelijkaardige nood bestaat binnen de mediasector: ook hier moet het materiaal zo uniform mogelijk geannoteerd worden, om nieuwe producten op basis van dit materiaal aan te kunnen maken, doorzoekbaarheid te garanderen, e.d. De studie leverde voor de deelnemers een aantal belangrijke inzichten op omtrent de verschillende aspecten van thesauri. Het eindrapport van het hele project schetst de grote lijnen en bevindingen.

56. Titel: Real-time Automated Media Production (2013-2014)

Financiering: iMinds MiX

Partners:

- a. KU Leuven (ITEC)
- b. VVMMA
- c. Small Town Heroes
- d. Televic Conference
- e. KU Leuven (CUO)
- f. UGent (IBCN)
- g. UGent (MMLab)

Beschrijving: Zowel omroep- als conferentiesystemen worden geconfronteerd met een stijgende vraag naar hoogwaardige, gemengde data-overlays tijdens het streamen van videocontent naar eindgebruikers. Er is nood aan een automatisch 'play-out'-systeem dat gebeurtenissen kan vastleggen en visualiseren, en dat in staat is om extra gegevens in real-time te integreren – terwijl een 'regisseur' (de DJ of conferentievoorzitter) slechts in geringe mate moet tussenbeide komen. Als onderdeel van dit project willen we niet alleen 'play-out'-systemen controleren; we zullen ook de controle van fysieke objecten in een studio onderzoeken, door middel van geautomatiseerde triggers – zoals het automatisch aanpassen van camera's en microfoons op basis van audio-, video- en contextuele informatie. Een goed semantisch begrip van de gestreamde content is daarvoor noodzakelijk, op basis van zowel contextuele informatie als de analyse van de inhoud. Bedoeling is om een semantisch play-out beslissingsondersteunend systeem te ontwerpen en te ontwikkelen. Dit project zal bijdragen tot 1) het creëren van een echt nieuw medium, 2) een efficiëntere en goedkopere postproductie en 3) het openstellen van TV Graphics naar het ecosysteem van Vlaamse web-agentschappen.

57. Titel: Ontwikkeling van een digitaal platform Nederlands leren (2013)

Financiering: EU – Europees Integratiefonds

Partners:

- a. KU Leuven (ITEC)
- b. KU Leuven (Centrum voor Taal en Onderwijs) coördinator
- c. KU Leuven (ILT)
- d. Taalblad.be
- e. UGent (MMLab)
- f. Universiteit Antwerpen
- g. Taaluniecentrum NVT
- h. VRT
- i. Vlaams Instituut voor Archivering (VIAA)
- j. Borgerhoff & Lamberigts tv
- k. Televic Education
- l. Theatergroep Fast Forward

Beschrijving: Het doel van dit project is een uitdagende en gevarieerde digitale omgeving te creëren voor alle anderstaligen die in Vlaanderen Nederlands willen leren en willen oefenen. Anderstaligen die in het Nederlands willen kunnen communiceren om in het dagelijks leven te kunnen functioneren, moeten de mogelijkheid krijgen om de talige vaardigheden die ze verwerven op vele manieren in te zetten en verder te ontwikkelen in een Nederlandstalige omgeving. Daarnaast wordt de inzet van ICT in het domein van het taalonderwijs algemeen beschouwd als een hefboom voor taalleren. Het ontwikkelen en gebruiken van een digitaal platform op maat van de leerder kan sterk complementair zijn met het NT2-onderwijs en met bestaande initiatieven in de taalstimulering in de vrije tijd. Dit wordt een platform dat anderstaligen, onafhankelijk of ze een NT2-cursus volgen of niet, de mogelijkheid wil bieden om op hun manier en wens aan de slag te gaan met Nederlands. De actualiteit vormt de inhoudelijke insteek. De projectuitvoerders worden bij de ontwikkeling en promotie van het platform ondersteund door een uitgebreide adviesgroep. Een gebruikersgroep, bestaande uit een diverse groep van anderstaligen, geeft feedback op inhoud en technologie zodat het platform motiverend, werkbaar en toegankelijk is voor elke gebruiker. De ontwikkeling van dit platform wordt mogelijk gemaakt door de financiële ondersteuning van het Europees Integratiefonds en het Vlaamse cofinanciering Integratiefonds.

Output: de gratis online leeromgeving Nedbox (www.nedbox.be)

58. Titel: Taggitus - Trainen en evalueren van een tagger om Latijnse teksten automatisch te annoteren (2016)

Financiering: contractonderzoek

Partners:

- a. KU Leuven (ITEC)
- b. Uitgeverij Brepols NV

Beschrijving: Dit project is een haalbaarheidsstudie, waarin we proberen na te gaan op welke manier de bestaande taaltechnologie geschikt is voor het automatisch annoteren van Latijnse teksten. Met annoteren bedoelen we hier part-of-speech (PoS) tagging en lemmatisering van woorden: d.w.z. zowel de woordsoort als het woordenboekwoord (lemma) wordt toegekend voor ieder woord in een tekst.

59. Titel: ASIMOV – Adaptive Semantics Induction for Multilingual matching in an enriched human resources environment (2016-2018). Zie 103

60. Titel: TECOL – Technology Enhanced Collaborative Learning (2016-2018)

Financiering: KU Leuven

Partners:

- a. KU Leuven (ITEC)
- b. Barco
- c. Televic

Beschrijving: Educatieve technologie wordt momenteel vooral ingezet voor de organisatie van het onderwijsgebeuren, de overdracht van informatie en het faciliteren van communicatie tussen actoren. Het leren zelf wordt in beperktere mate technologisch ondersteund, waardoor er een duidelijke nood is aan oplossingen die interactief en collaboratief leren meer centraal plaatsen. Deze nood leidde tot de uitbouw van het TECOL-project, om recente educatief technologische oplossingen structureel binnen te brengen in het contactonderwijs, zowel in de hoorcolleges als in de werksessies en practica als tussen verschillende locaties. We zetten hiervoor de nodige proeftuinen op die we op hun effectiviteit evalueren.

Output:

- Inrichting van learning spaces aan de KU Leuven campus Kulak Kortrijk, waaronder 4 collaboratieve leerruimtes onder andere in het Edulab, een connected classroom om remote groepen met elkaar te verbinden, een virtual classroom om individuele studenten actief te laten participeren in lessen en ruimtes voor domein specifieke vaardigheden, zoals interactieve taallabo's. Daarnaast werd ook het open

leercentrum ingericht met technologie-ondersteunde groepsruimtes en een presentatieruimte.

- Verzamelen van didactische best practices in de verschillende leeromgevingen (Interactief hoorcollege, collaboratieve leerruimtes, remote classroom en virtual classroom).

Outreach: (zie ook <https://www.kuleuven-kulak.be/tecol/>)

- We ontvingen heel wat verschillende types kennisinstellingen in ons Edulab met als doelstelling kennismaking met andere instituten en het delen van ervaringen met technologie ondersteund collaboratief contactonderwijs:
 - o Business schools vb. Vlerick, IESE Barcelona, Cambridge Judge Business School, EDHEC
 - o Universiteiten – zowel nationaal als internationaal vb. Université de Mons, UCLouvain, UCLondon, Imperial College London, RU Nijmegen, Eindhoven, UVA, Breda, Leiden, Rotterdam, Tilburg, Hulskamp, Erasmus Rotterdam, UCC Kopenhagen, London Metropolitan University, City University, University of Greenwich, University of Bath, Cardiff University, University of Reading, University of East London, Toulouse, Rouen, Oulu Finland, UCLL
 - o Hogescholen – zowel nationaal als internationaal vb. Vives Brugge, Thomas More, HVA, Fontys, Haagse Hogeschool, Nederlandse Politieacademie
 - o Secundaire scholen: vb. RHIZO Lyceum OLV Vlaanderen, Sint Lodewijkscollege Brugge, Heilig Hart College Waregem, Prizma groep Izegem, Dienst Pedagogisch Begeleiders West-Vlaanderen
- TEA-CALL (2016): workshop gericht op de docenten van de Kulak – docenten laten de nieuwe beschikbare mogelijkheden van het TECOL project uittesten en brainstorm over de mogelijke toepassingen en het gebruik.
- Kortrijk wordt de aula van Kulak - Raes A., & Depaep F. (2016). "Smart Education" Hoe kan technologie kwaliteitsvol onderwijs ondersteunen?
- TEA-CALL (2017): workshop gericht op de docenten van de Kulak – hands-on training geven over de presentatieruimte in de bibliotheek en voorstelling onderzoeksdesigns.
- Workshop op een studiedag in Maastricht: het gebruik van e-tools in het onderwijs (2017).
- workshop op de onderwijsdag FEB KULeuven (2017)
- Junior college (2017): Raes A., & Depaep F. (2018). De digitale klas: Utopie of werkelijkheid?
- workshop op de onderwijsdag RU Nijmegen rond best practices in de collaboratieve leerruimte (2018).
- TEA-CALL (2018): workshop gericht op de docenten van de Kulak – de beschikbare mogelijkheden van zowel de remote als de virtual classroom demonstreren en docenten laten uittesten, aangevuld met een brainstorm over de didactische voorwaarden voor een succesvol multilocatieleren.
- imec Roadshow – demonstratie van een virtual classroom (2018).
- workshop op de Onderwijsdag van de Arteveldehogeschool (2018).
- studiedag aan het einde van het project, met als doelstelling (1) De belangrijkste bevindingen en good practices binnen het project te delen met het ruimer

onderwijslandschap, (2) hands-on workshops aan te bieden om aan de slag te gaan met de TECOL-technologie en (3) samen in gesprek te gaan over de toekomst van digitalisering en technologie binnen onderwijs.

- Het project en de bijhorende resultaten werden gepresenteerd in talrijke media (voor een overzicht; zie <https://www.kuleuven-kulak.be/tecol/#nieuws>):
 - o VTM-nieuws
 - o Focus|WTV nieuws
 - o Radio 2
 - o Krant van West-Vlaanderen
 - o De Tijd
 - o Edu&Care
 - o Trends
 - o Made in West-Vlaanderen
 - o Het Laatste Nieuws
 - o KanaalZ
 - o Het Nieuwsblad
- De resultaten werden daarnaast ook gepresenteerd op verschillende conferenties.

61. Titel: VALILEX – Theoretical and empirical validation of lexical competence in French within the Common European Framework of Reference (2014-2018)

Financiering: KU Leuven – BOF

Partners:

- a. KU Leuven
- b. Universiteit Amsterdam

Beschrijving: Het Europees Referentiekader voor Talen (ERK) is wellicht het belangrijkste beleidsdocument voor het vreemdetalenonderwijs in Europa. Ondanks zijn populariteit zijn de taalspecifieke invullingen hiervan theoretisch en empirisch onvoldoende gevalideerd. Dit project wil een bijdrage leveren aan een theoretische en corpusgebaseerde empirische validering van het ERK, die tegelijkertijd rekening houdt met bestaande taalvaardigheidstheorieën. Concreet willen we bepalen wat de gemeenschappelijk woordenschat van moedertaalsprekers is voor het uitvoeren van de B1-taalhandelingen beschreven in het ERK en willen we vaststellen wat de minimum woordenschat is die vreemdetaallearners nodig hebben om luister- en spreektaken op dat niveau succesvol te kunnen uitvoeren.

Outreach: Met dit project zullen we een theoretische en empirische bijdrage leveren aan het ERK en aan het domein van de vreemdetaalverwerving. De resultaten die uit het project voortvloeien, zullen impact hebben op zowel het Europese als nationale onderwijsbeleid, op handboeken voor het taalonderwijs (uitgeverijen) en op alles wat met taaltoetsing te maken heeft.

62. Titel: Enhancing audio-visual input in order to foster second language learners depth of lexical knowledge (2014-2018)

Financiering: FWO – postdoctoraal mandaat

Partners: KU Leuven (ITEC)

Beschrijving: In dit project onderzoeken we de rol van drie technieken om woordenschatverwerving via video bij vreemdetallearders Frans te stimuleren. In een eerste studie onderzoeken we de rol van verschillende types ondertitels (in de moedertaal, in de doeltaal, met kernwoorden, etc.) voor woordenschatverwerving. In studie 2 onderzoeken we het effect van het herhaaldelijk presenteren van nieuwe woorden (frequentie van woordenschat) in video's. In studie 3 bestuderen we het effect van het vooraf (voor het bekijken van de video) aanleren van nieuwe woordenschat. Om woordenschatverwerving te meten, gebruiken we zowel interviews en traditionele pen/papier testen als eye-tracking en reactietijdmetingen.

Outreach: Aangezien de resultaten van dit onderzoek direct relevant zijn voor leerkrachten vreemdetalenonderwijs werd ingezet op navormingen voor leerkrachten vreemdetalen rond het gebruik van video voor woordenschatverwerving. Ondertussen hebben meer dan 450 leerkrachten vreemdetalen een sessie hierrond gevolgd op een van onderstaande initiatieven:

- o Montero Perez, M. (2018). TV-kijken als bron van woordenschatverwerving. Studiedag Proff, Gent, 9 maart 2018.
- o Montero Perez, M. (2018). TV-kijken en woordenschat. Dag van Frans, Kortrijk, 5 Februari 2018.
- o Montero Perez, M. (2017). Woordenschatverwerving in een vreemde taal. Didactiekdag ANBF, Brussel, 17 mei 2017.
- o Montero Perez, M. (2017). TV-kijken en woordenschat. AVL Nascholing Lezen of TV-kijken als bron voor woordenschatverwerving, Leuven, 26 April 2017.
- o Montero Perez, M. (2016). TV-kijken en woordenschatverwerving. Studieavond: De woorden kwijt / Lost for words. Gent, 21 April 2016.

63. Titel: LEAPS - LEarning analytics for AdaPtive Support (2016-2018)

Financiering: imec.ICON

Partners:

- a. KU Leuven (ITEC) – coördinator
- b. UGent (IDLab)
- c. KU Leuven (mintlab)
- d. Sensotec

- e. BLCC
- f. Uitgeverij Pelckmans
- g. Basisschool De Ring Roeselare
- h. Basisschool De Plataan Roeselare
- i. Basisschool Bellevue Izegem
- j. Basisschool Het Open Groene Marke
- k. gvbs Sint Rafaëlsschool Izegem
- l. vrije basisschool Centrumschool Ingelmunster
- m. gvbs Sint-Vincentius Lendeledede
- n. Stedelijke basisschool Noord Hulste
- o. gemeentelijke Boudewijnschool (Pienter) Kuurne
- p. gemeentelijke basisschool Zuid Harelbeke.

Beschrijving: Het LEAPS-project werd opgestart om een zelflerend analysesysteem te ontwikkelen voor aangepaste leerondersteuning. Een systeem dat geïntegreerd kan worden in hulpsoftware voor mensen met leesproblemen, professionele communicatie en educatieve spelletjes. Het consortium bestaat uit partners met uiteenlopende expertise, onder meer in statistische, computer-, onderwijs- en cognitieve wetenschappen, mens-computerinteractie, softwareontwikkeling en educatieve publicaties. Dankzij deze interdisciplinaire samenwerking komt een commercialiseerbare, maar vooral broodnodige oplossing binnen handbereik.

64. Titel: Language learning through task-based interaction with chatbots (2014-2018)

Financiering: doctoraatsbeurs

Partners:

- a. KU Leuven (ITEC)
- b. UCL (CENTAL)

Beschrijving: Considering the existing evidence of the effectiveness of synchronous computer-mediated communication on the development of L2 oral proficiency (Lin, 2015), combined with the scarcity of opportunities to interact with native speakers in most foreign language learning contexts, ICALL researchers have long been interested in developing systems that would allow a learner to have a meaning-focused interaction in L2 with a conversational agent (e.g. Underwood, 1982; Wachowicz & Scott, 1999; Fryer & Carpenter, 2006). During the last three decades, various dialogue systems have been developed as proofs of concept, using different natural language processing techniques to accept unconstrained learner production, either spoken (see Eskenazi, 2009 for an overview) or written (e.g. Price, Bunt, & McCalla, 1999; Wolska & Wilske, 2011). In the conceptualization of those conversational systems, a computer-guided approach of dialogue has proven both more effective and more feasible than the open-ended one commonly seen in chatbots (e.g. Petersen, 2010).

65. Titel: Upscaling the Geo-temporal Enrichment, exploration and exploitation of Scientific Collections (UGESCO) (2016-2019)

Financiering: Belspo BRAIN-be

Partners:

- a. KU Leuven (ITEC)
- b. UGent (IDLab) – coördinator
- c. UGent (Cartogis)
- d. ULB
- e. Cegesoma

Beschrijving: Het UGESCO-project beoogt de ontwikkeling van tools voor geo-temporele (meta)data-extractie en verrijking zodat de bestaande collectie-items (i.c. fotoarchieven) met elkaar gekoppeld kunnen worden en met extra informatie kunnen worden uitgebreid, zodat spatio-temporele collectie mapping voor interactieve opvragingen wordt vergemakkelijkt. Met het oog op kwaliteitsverbetering en validatie zal ook uitgetest worden hoe om te gaan met microtaskcrowdsourcing, metadatamanagement en filtertechnieken. Een aangepaste gebruikersinterface zal worden aangeboden zodat wetenschappelijke interpretatie op basis van de gelinkte data sets wordt vergemakkelijkt.

66. Titel: De predictie van de receptieve lexicale competentie van taalleerders via een gepersonaliseerd en adaptatief model gebaseerd op een betekenis-gebaseerde aanpak

Financiering: Aspirant F.R.S.-FNRS

Partners:

- a. KU Leuven (ITEC)
- b. UCL (CENTAL)

Beschrijving: Dit project, gesitueerd in het computerondersteund taalleren en de taaltechnologie, heeft als doel de receptieve lexicale vaardigheid van vreemdetaallearners van Frans en Nederlands in een leesactiviteit automatisch te voorspellen. Aan de ene kant willen we het receptieve lexicon van een taalleerder empirisch evalueren via een betekenis-gebaseerde aanpak. We baseren ons op een corpus waar de moeilijkheid van elke woordbetekenis geannoteerd is door een groep vreemdetaallearners. Aan de andere kant willen we de impact van persoonlijke kenmerken op de voorspelling van woordenschatkennis bestuderen. We ontwikkelen daarom een computationeel model die deze kenmerken integreert en zich aanpast aan de individuele leerder.

Output:

- NT2Lex een geïntegreerde lexicale database voor Nederlands vreemde taal (tool + API): <http://cental.uclouvain.be/cefrlex/> (binnenkort beschikbaar)

- Data en statistische modellen voor het voorspellen van lexicale moeilijkheden in Frans en Nederlands (tool + API): <http://cental.uclouvain.be/predicomplex/> (toekomstplannen)

Outreach: De resources die we ontwikkelen in dit project zullen voor niet-commercieel gebruik ter beschikking gesteld worden en kunnen dus ten eerste van nut zijn voor verder onderzoek. Verder proberen we die ook gebruiksvriendelijk te maken voor een groter publiek en met name taalleerkrachten en -leerders, die we via presentaties in scholen en andere pedagogische centra kunnen bereiken.

67. Titel: Impact of form- or meaning-focused instruction on L2 lexical acquisition: A behavioural and neurological study (2016-2019)

Financiering: FWO (onderzoeksproject)

Partners:

- a. KU Leuven (ITEC)
- b. KU Leuven (Faculteit Psychologie & Pedagogische Wetenschappen)

Beschrijving: De efficiëntie van vorm- of betekenisgerichte approaches is een centraal vraagstuk in het onderzoek rond vreemdetaalverwerking. In dit project wordt de impact van drie vormen van vreemdetaalinstructie (vormgerichte instructie in context, vormgerichte instructie zonder context, en geen vormgerichte instructie in context) bestudeerd voor woordenschatverwerving in een vreemde taal. Om woordenschatverwerving te meten zullen drie meetinstrumenten met toenemende graad van gevoeligheid worden ingezet: expliciete tests (vb. traditionele pen/papier testen), impliciete tests (i.e., het meten van reactietijden) en neurologische tests (registreren van Event-related potentials, ERPs). Twee experimentele studies zullen worden afgenomen bij 16/17-jarige leerders Frans (N=300).

Outreach: Dit project zal concrete implicaties hebben voor het aanleren van woordenschat in de klaspraktijk. Daarom zal sterk ingezet worden op navormingsinitiatieven voor leerkrachten vreemdetaal. Deze inzichten zullen implicaties hebben voor ontwikkelaars van handboeken en lesmaterialen.

68. Titel: FanTALES - Fanfiction for the Teaching and Application of Languages through E-Stories (2017-2020)

Financiering: EU – Erasmus+

Partners:

- a. KU Leuven (coördinator)
- b. Malmö University
- c. Goethe Universität Frankfurt
- d. Georg-Forster-Gesamtschule Woerrstadt
- e. Lunds dans- och musikalgymsnasium
- f. Guldensporencollege Kortrijk
- g. Sint-Jozefscollege Aalst

Beschrijving: FanTALES stands for the creation of multilingual interactive fan stories in schools for a more inclusive world. Stories can help foster empathy and understanding of context and culture, which are critical in a plurilingual and pluricultural world. In the digital wilds, communities of practice can be encountered where stories are told, retold, and experienced interactively. Key examples of such practices are fan fiction writing, a form of creative writing that transforms popular media in some way, and the development of interactive fiction, a form of non-linear narrative that verges on the world of gaming. While these practices — in particular the dialogic interaction with stories that is inherent in them — can effectively eliminate cultural stereotypes and discrimination, thereby fitting in well with the current educational agenda of Europe that promotes tolerance and nondiscrimination, fan fiction and interactive fiction are only slowly finding their way into language education.

B. UNIVERSITEIT ANTWERPEN – OVERZICHT PROJECTEN

69. Titel: Artificiële Creativiteit in visuele communicatie en kunst: een algoritme voor spitsvondige, evoluerende conceptontwikkeling en datavisualisatie. (2008-2012)

Financiering: IOF doctoraat

Partners:

- a. UAntwerpen (CLiPS)
- b. Sint-Lucas Antwerpen (EMRG)

Beschrijving: Met behulp van technieken uit de Artificiële Intelligentie wordt een computeralgoritme ontwikkeld dat een inhoudelijke opdracht (of dataset) ontleedt op kernbegrippen en relaties, daarbij relevante informatie opzoekt, verwerkt, bundelt en deze tenslotte in relatie brengt tot creatieve en visuele oplossingen. Dit algoritme wil menselijke creativiteit nabootsen door bestaande concepten te verbinden en te hercombineren tot er innovatieve visuele uitvoer uit opborrelt. De visuele uitvoer evolueert mee naargelang de inhoudelijke data wijzigt en uitbreidt.

Output: Python Pattern software module, artwork

Outreach: Expositions, IMEC building wall art

70. Titel: ALADIN: Adaptation and Learning for Assistive Domestic Vocal Interfaces (2011-2014).
Zie hierboven

71. Titel: deLearyous: training van Interpersoonlijke communicatie door natuurlijke taalinteractie met autonome virtuele karakters (2010-2012)

Financiering: VLAIO (IWT) Tetra

Partners:

- a. UAntwerpen (CLiPS)
- b. KU Leuven (Groep T)

Beschrijving: Het doel van het deLearyous-onderzoeksproject is het ontwikkelen van een interactieve serious 3D-game voor het trainen van interpersoonlijke communicatievaardigheden binnen een professionele context, bvb. in een werkgever-werknemer- of klant-bediende-relatie. Het spel bestaat erin de trainee te laten interageren met autonome virtuele karakters, die op een realistische en expressieve manier inspelen op de input van de speler. Op deze wijze kan de trainee verschillende gedragspatronen en rollen

op eigen tempo inoefenen, en dit in een veilige, virtuele omgeving. De rol van CLiPS in het project is de ontwikkeling van algoritmen en methodes voor emotie-analyse van tekst, het detecteren van onderwerpen in tekst, en dialoogvoering.

Output: doctoraat, demo's

72. Titel: Een veiliger internet: het (semi)automatisch herkennen van internetpedofilie in meertalige online social networks (2010-2013)

Financiering: BOF doctoraat

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (Trics)

Beschrijving: In dit project stellen we enerzijds een methodologie voor om de manuele controle van peer-to-peernetwerken te (semi)automatiseren en anderzijds een methodologie voor de automatische extractie en analyse van stijkenmerken (geassocieerd met persoonlijkheid, leeftijdsgroep en misleidend taalgebruik) die we willen toepassen op individuele pedofielen en groepen van pedofielen in chatruimtes.

73. Titel: TTNWW: TST Tools voor het Nederlands als Webservices in een Workflow (2010-2012).
Zie hierboven

74. Titel: Stylene, een webservice voor stylometrie- en leesbaarheidsonderzoek voor het Nederlands (2010-2012).

Financiering: EWI

Partners:

- a. UAntwerpen (CLiPS) - coördinatie
- b. UGent (LT3)

Beschrijving: Doel van het project is de implementatie van een robuust, modulier systeem voor stylometrie- en leesbaarheidsonderzoek op basis van bestaande technieken voor automatische tekstanalyse en zelflerende technieken, en de ontwikkeling van een web service die onderzoekers in de HSS toelaat teksten te analyseren met behulp van het systeem. Op die manier wil het project recente vooruitgang op het gebied van het

computationeel modelleren van stijl en leesbaarheid beschikbaar maken voor onderzoek in de sociale en geesteswetenschappen.

Output: demo's, software

75. Titel: Een publiek beschikbare Economische Onzekerheids-Indicator voor de G8 landen m.b.v. tekst mining (2014-2015)

Financiering: Nationale Bank van België

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (Applied Data Mining)

Beschrijving: In dit project focussen we op de vraag: hoe kunnen we onzekerheid omtrent economisch beleid meten? Recent ontwikkelden we een EPU (Economic Policy Uncertainty) index voor België, door online nieuwsartikels van de grootste kranten te minen. Gegeven de veelbelovende resultaten die we bekwamen, wensen we deze tekst mininggebaseerde methodologie toe te passen voor andere landen (de G8 landen) en zullen we een publieke website creëren waar deze index wekelijks automatisch zal worden geüpdatet.

Outreach: website met indexen

76. Titel: PARIS: Personalised AdveRtisements built from web Sources (Gepersonaliseerde advertenties gebouwd uit webbronnen) (2012-2016). Zie [hierboven](#)

77. Titel: Een middeleeuws Styloom? Een verkenning van de Universele Styloom-Hypothese in middeleeuws proza. (2012-2015)

Financiering: FWO Postdoc

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (ACDC)

Beschrijving: Computergestuurd onderzoek binnen de kwantitatieve stijlleur of stylometrie heeft geleid tot de "Styloom-hypothese": de niet-bewezen aanname dat iedere auteur zo een individuele schrijfstijl heeft dat dit "stilistisch DNA" kwantitatief kan gemodelleerd worden. In dit project wordt casusgewijs getoetst in welke mate deze aanname opgaat voor

middeleeuwse literatuur. Dat is uitdagend, want middeleeuwse teksten bevatten veel variatie (bv. spelling) die niet teruggaan op de oorspronkelijke auteur van teksten, maar op latere kopiïsten. Dit onderzoeksproject legt de nadruk op twee casussen. Ten eerste wordt de stylometrie toegepast op Brabantse geestelijke teksten uit de dertiende en veertiende eeuw (o.m. Hadewijch en Jan van Ruusbroec). Ten tweede zal de stijl bestudeerd worden van Latijnse literatuur uit de elfde eeuw van Anglo-Vlaamse herkomst (o.m. Drogo, Goscelinus).

78. Titel: Abstracte regels of statistisch leren? De impact van lexicale en sublexicale homofonie tijdens het spellen en lezen van homofone werkwoordvormen. (2012-2015)

Financiering: FWO Project

Partners: UAntwerpen (CLiPS)

Beschrijving: Homophone intrusions in the spelling of regularly inflected Dutch verb forms are used to address a central question in psycholinguistics – and cognitive science in general: do people rely on symbolic mental rules or on a knowledge base that captures the co-occurrence probabilities in the learning domain (statistical learning)? Earlier findings in our research group indicated an effect of homophone dominance in the pattern of intrusion errors when spelling homophonic verb forms: such errors occur more often when the target is the lower-frequency homophone and the intruder the higher-frequency homophone. This is compatible with a statistical learning view but cannot reject a rule-based account enriched with a frequency-sensitive mechanism. To disentangle the two accounts we will compare error patterns in the lexical and sublexical domains. An effect of homophone dominance at the sublexical level cannot be explained by a rule model. Errors in the lexical and sublexical domains will be studied in spelling and reading tasks. Finally, we will attempt to simulate the experimental patterns with two types of computational models: a symbolic model, using morphemes and rules, and a memory-based model, storing whole word forms only and using a similarity metric that can 'discover' patterns in its memory store. Together, the experimental and simulation data should enable us to formulate an answer to the question about mental rules.

79. Titel: Analyseren van de impact van nieuws- en beursberichten op Belgische aandeleprijs via tekst mining (2011-2013)

Financiering: Nationale Bank van België

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (Applied Data Mining)

Beschrijving: In dit project zullen we nagaan hoe algemene en beursspecifieke nieuwsberichten met geavanceerde textminingtechnieken gebruikt kunnen worden voor het automatisch voorspellen van het effect op Belgische aandelenprijzen. Inzichten zullen bekomen worden in welke nieuwsleveranciers en welke combinaties van woorden het grootste effect hebben. Het ontworpen systeem zal geëvalueerd worden als trading tool, alsook als beslissingsondersteunend systeem voor investeerders.

80. Titel: De computationele leerbaarheid van morfologisch complexe talen. (2006-2012)

Financiering: FWO Postdoc

Partners: UAntwerpen (CLiPS)

Beschrijving: Traditionele systemen voor spellingcontrole maken gebruik van een woordenlijst. Wanneer een woord niet voorkomt in de woordenlijst, markeert het systeem het woord als "fout". Recente systemen benaderen het probleem van spellingcontrole voor agglutinerende talen vanuit een andere invalshoek: een woord wordt beschouwd als een spelfout, als het niet kan worden gegenereerd door een achterliggend morfotactisch model. In dit project onderzoeken we hoe een spellingchecker kan gebruikt worden als hulpmiddel bij het automatisch leren van een morfotactisch systeem voor het Swahili.

81. Titel: Datafusie en gestructureerd input- en output- machinaal leren technieken voor geautomatiseerde klinische codering (2014-2017).

Financiering: VLAIO (IWT) doctoraatsbeurs

Partners: UAntwerpen (CLiPS)

Beschrijving: This project will improve the state of the art in automated clinical coding by analyzing heterogeneous data sources and defining them in a semantic structure and by developing novel data fusion and machine learning techniques for structured input and output.

82. Titel: De rol van de semantiek in het modelleren van het bilinguale mentale lexicon. (2016-2019)

Financiering: FWO aspirant

Partners: UAntwerpen (CLiPS)

Beschrijving: Bilingualen, mensen die tegelijk twee of meer talen kennen en gebruiken, zijn een interessante bron van informatie voor het ontdekken van de interne structuur van het taalsysteem. De huidige modellen voor bilinguaal lezen van woorden verklaren de meeste empirische gegevens, maar tonen weinig aandacht voor de rol van betekenis. Ze modelleren woorden bovendien in isolatie. In dit project ontwikkelen we computermodellen van zinsverwerking waardoor betekenis een centrale rol speelt die het verdient en nieuwe verklaringen gegeven kunnen worden voor verschillende fenomenen in de bilinguale verwerking van woorden.

83. Titel: Diepe taalkundige features voor computationele stylometrie. (2015-2018)

Financiering: FWO aspirant

Partners: UAntwerpen (CLiPS)

Beschrijving: In dit project trachten we nieuwe, meer linguïstische tekstkenmerken te ontdekken die kunnen helpen in het stylometrisch onderzoek naar auteursprofilering. Meer specifiek gaan we op zoek naar tekstkenmerken gerelateerd aan semantiek en discourse, zoals bijvoorbeeld semantische densiteit. Onze hypothese stelt dat deze nieuwe tekstkenmerken informatie bevatten over de auteur, aanvullend aan oppervlakkige features, zoals bag-of-words, die momenteel vaak gebruikt worden in machine learning experimenten. Deze kenmerken worden getest in voornamelijk 3 use cases: herkenning van leeftijd, geslacht en persoonlijkheid van de auteur.

84. Titel: Automatic Compound Processing (Automatisch verwerken van samenstellingen) (2012-2013)

Financiering: NTU

Partners:

- a. UAntwerpen (CLiPS)
- b. Universiteit van Tilburg (ILK)
- c. Universiteit Potchefstroom

Beschrijving: The central problem to be addressed in this project concerns a multidisciplinary investigation into sharing of knowledge and resources between closely-related languages, specifically relating to the automatic processing of compounds. Specifically, we will explore the possibility to create new knowledge about closely-related languages, and efficiently develop additional, more advanced resources for (a) compound segmentation; and (b) the semantic analysis of compounds.

Output:

- software voor de automatische analyse van samenstellingen in het Nederlands en Afrikaans
- annotatierichtlijnen voor de analyse van samenstellingen
- datasets
- een workshop

85. Titel: AMiCA: Automatic Monitoring for Cyberspace Applications (2013-2016)

Financiering: IWT-SBO

Partners:

- UAntwerpen (CLiPS) - coördinator
- UGent (LT3)
- UGent (IBCN)
- UAntwerpen (MIOS)
- KU Leuven (PSI)

Beschrijving: Hoewel het gebruik van het internet door kinderen meestal volkomen veilig en verrijkend is, zijn er ook risico's verbonden aan het gebruik van vooral social networking sites (SNS). Verschillende initiatieven werden opgezet om kinderen te beschermen tegen deze potentiële bedreigingen door preventie en door in te grijpen als het misloopt, maar gezien de enorme overvloed aan informatie is het onmogelijk voor de belanghebbenden om alle potentieel schadelijke situaties (zoals cyberpesten, suïcidaal gedrag of grooming door pedofielen) te traceren. Het automatisch traceren van schadelijke inhoud en schadelijke gedragingen vereist een crossmedia filtering aanpak, die onmiddellijk risico's kan detecteren, die kritieke situaties doorgeeft aan moderatoren van sociale netwerksites en die accurate kwantitatieve informatie verstrekt zodat longitudinale en beleidsondersteunende informatie ter beschikking komt over de online veiligheid van kinderen. Deze filtering zal gebeuren rekening houdend met de verwachtingen en online ervaringen van adolescenten en kinderen en ook gebruik makend van een aanpak die op een respectvolle en zorgzame manier omgaat met privacygevoelige informatie.

Output:

- sociale media taalnormaliseringssoftware
- corpora voor profilering
- software voor profilering uit tekst
- sentimentanalysesoftware
- Keyword spotter voor "haatspraak"

Outreach: Zie ook <http://www.amicaproject.be/outreach>

- AMiCA genomineerd voor een LIA-award

- Radio-interview met Kathleen Van Royen: “Femme de la rue”, maar dan online (Radio 1, 17/11/2015)
- Media:
 - o Jongeren willen dat Facebook sneller ingrijpt bij aanstootgevende inhoud (De Morgen en De Standaard – 19/3/2015)
 - o Aandacht voor onderzoek naar cyberpedofilie in “De Ideale Wereld” (17/03/2014)
 - o Taalgebruik verradt pedofiel (De Standaard, 17/03/2014)
 - o De taalcomputer leest wie u bent (De Standaard Taalbijlage, 06/11/2013)
 - o Klasse confronteert met cyberpesten (VTM Nieuws, 28/05/2013)
 - o Antwerpse Hoeren-facebookpagina: 11% jongeren slachtoffer van cyberpesten (Knack.be – 03/1/2013)
 - o Meer dan helft jongeren reageert niet op cyberpesten (De Standaard online, 20/11/2012)
 - o Taalcomputer ontmaskert pedofielen in chatrooms (De Morgen, 15/12/2011)

86. Titel: Webservices voor tekstanalyse: profilering en opiniedetectie (2014-2015)

Financiering: IOF Proof of Concept

Partners:

- a. UAntwerpen (CLiPS)
- b. Allzenz
- c. X-trade
- d. STIL

Beschrijving: We willen commerciële webservices opzetten die opinies en auteursprofielen (leeftijd, geslacht, persoonlijkheid, opleidingsniveau, regio) uit tekst winnen, op basis van automatische tekstanalyse. In dit project ontwikkelen we de basistechnologie: data verzamelen en annoteren, machine learning algoritmen trainen en een serverplatform opzetten. In een vervoltraject richten we een spin-off van de onderzoeksgroep op. De voorgestelde taaltechnologie is nuttig voor een waaier van big data toepassingen, en bestaat nog niet voor het Nederlands en slechts gedeeltelijk voor het Engels.

Output/outreach: Spin-off bedrijf Textgain. Zie <http://www.textgain.com>

87. Titel: Bootstrapping operaties in de taalverwerving: een computationeel psycholinguïstische benadering. (2014-2017)

Financiering: BOF TOP

Partners: UAntwerpen (CLiPS)

Beschrijving: We onderzoeken hoe kinderen abstracte linguïstische categorieën leren. Daarvoor worden computersimulaties van bootstrapping operaties ontworpen: hoe kan kennis uit één domein gebruikt worden om kennis te verwerven van een ander domein? In de simulaties wordt reële omgevingstaal aan jonge kinderen gebruikt en wordt nagegaan hoe grammaticale woordcategorieën en grammaticaal geslacht via een combinatie van distributionele, fonologische en morfologische bootstrapping verworven kunnen worden.

Output: computationele modellen van taalverwerving

88. Titel: De interactie van gender en sociale klasse in de chattaal van Vlaamse tieners (2015-2018)

Financiering: FWO-project

Partners: UAntwerpen (CLiPS)

Beschrijving: Social class differences in teenage speech remain largely unexplored, while gender has been focused on in quite a lot of sociolinguistic research on adolescent peer group language. The interest in gender differences has also pervaded the research on informal computer-mediated communication (CMC) and more specifically on the online writing practices of adolescents in chat or texting media, but then again, the link with social class is generally absent. Yet some studies (though not on CMC) suggest that gender differences manifest themselves in different ways in different social class groups. The present research is a first attempt to fill this gap, by focusing on the interaction between social class and gender in Flemish chat language produced by adolescents with a low versus a high level of education. In a computational component, we will investigate the prediction of social class and gender from text.

Output: software voor predictie van opleidingstype, sociale klassen en gender uit tekst

89. Titel: ACCUMULATE: ACquiring CrUcial Medical information Using LAnguage Technology (2016-2019). Zie boven

90. Titel: INSIGHT: Intelligent Neural Systems as InteGrated Heritage Tools (2016-2021).

Financiering: Belpo – BRAIN-be (POD wetenschapsbeleid)

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (ACDC) (coördinatie)
- c. ULg
- d. UGent

Beschrijving: Het INSIGHT-project zet in op de toepassing van geautomatiseerde algoritmes uit het domein van de artificiële intelligentie om culturele erfgoedinstellingen te ondersteunen bij hun lopende digitalisatie-initiatieven en het beheer van hun snel groeiende digitale collecties. We focussen op de recente vooruitgang in machine-gebaseerd leren, waar de toepassing van diepe neurale netwerken tot belangrijke doorbraken heeft geleid in natuurlijke taalverwerking en computationele perceptie-studies. We gaan na hoe geavanceerde algoritmes kunnen gebruikt worden om op (semi-automatische wijze) digitale objecten te beschrijven en te catalogeren, in het bijzonder deze waarvoor geen of slechts nauwelijks metadata beschikbaar zijn. Dit project plaatst de collecties van twee federale museumclusters te Brussel centraal (Koninklijke Musea voor Schone Kunsten en Koninklijke Musea voor Kunst en Geschiedenis) en neemt zich voor hun digitale collecties klaar te maken om opgenomen te worden in Europeana.

Output: Deep neural network software voor beeldverwerking en tekstanalyse

91. Titel: Optimalisatie van het aanpassingsvermogen van klinische informatie-extractiesystemen: deep learning en user-feedback propagatie-technieken (2017-2021)

Financiering: VLAIO-Baekeland

Partners:

- a. UAntwerpen (CLiPS)
- b. Lynxcare

Beschrijving: Tegenwoordig zijn grote hoeveelheden niet-gestructureerde tekstuele medische gegevens beschikbaar (bijvoorbeeld klinische rapporten), wat mogelijkheden opent voor optimalisatie van de kwaliteit van de gezondheidszorg en voor de veiligheid van de patiënten. Hoewel de taaltechnologie al goede instrumenten en systemen beschikbaar heeft voor de automatisering van de verwerking van deze medische documenten, schiet de accuraatheid ervan dikwijls tekort wanneer de context van de informatie-extractie wijzigt (andere medische specialiteit, ander ziekenhuis, andere specialist). Dit project bestudeert de mogelijkheid van een schaalbare aanpak die zich aanpast aan dergelijke nieuwe contexten. Om dit doel te bereiken zullen we aanpakken gebaseerd op diepe neurale netwerken, het "human-in-the-loop" paradigma en persistent leren onderzoeken. Dit project gebeurt in samenwerking met Lynxcare Clinical Informatics, een medisch IT-bedrijf dat zich

concentreert op het verbeteren van toegang tot medische informatie en reductie van administratieve kosten in ziekenhuizen.

Output: verwacht: software voor klinische tekstanalyse

92. Titel: De invloed van politiek nieuws op burgers, en van burgers op het politieke nieuws, in tijden van sociale media. Theoretische uitdagingen en empirische mogelijkheden. (2017-2021)

Financiering: BOF-GOA

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (Departement Politieke wetenschappen)
- c. UAntwerpen (Departement Communicatiewetenschappen)

Beschrijving: In een democratie hebben burgers nood aan politieke kennis. De massamedia worden traditioneel aanzien als een belangrijke bron om burgers van politieke informatie te voorzien. Studies over agenda-setting en framing hebben veelvuldig aangetoond dat de nieuwsmedia een duidelijke invloed hebben op wat mensen weten, en hoe ze over politiek denken. De vraag is in welke mate deze klassieke inzichten overeind blijven in het digitale tijdperk. Het stijgend belang van het internet en de sociale media in het bijzonder als kanalen van informatie en communicatie hebben de manier waarop mensen leren over de wereld, en de politiek in het bijzonder, drastisch veranderd. Zo staat bijvoorbeeld de rol van de media als agenda-setter en 'framer' onder druk, omdat sociale media het causale verband, van massamedia naar publiek, ter discussie stellen. Journalisten worden meer en meer beïnvloed door discussies op blogs, Facebook, Twitter en andere platformen. Bovendien hebben politici meer mogelijkheden om de traditionele media te omzeilen en het publiek direct te beïnvloeden. Kortom, we bestuderen hoe mensen nieuws consumeren en hoe ze daardoor beïnvloed worden, maar ook hoe journalisten en politici op hun beurt beïnvloed worden door de reacties van mensen op het nieuws. Digitale media stellen niet enkel de klassieke theoretische inzichten in vraag, ze bieden tegelijk nieuwe mogelijkheden om informatiestromen en hoe mensen ermee omgaan te bestuderen. Vandaag is het mogelijk om agenda-setting en framing meer accuraat te bestuderen, gebaseerd op grotere databestanden van mediaboodschappen en de publieke reacties daarop (commentaren, likes, tweets). Bovendien laat het analyseren van digitale teksten toe om demografische profielen en opinies van burgers in kaart te brengen en onze kennis van de modererende factoren van agenda-setting en framing effecten te bestuderen.

Om deze ongeziene bron van geschreven taal en digitaal gedrag te interpreteren, kiezen we voor een multidisciplinaire samenwerking tussen computerlinguïstiek, data mining en sociale wetenschappen. De agenda-setting en framing theorie worden in een digitale context getoetst aan de hand van 'big data' analyses. Computerlinguïstische technieken zullen worden gebruikt om automatisch de onderwerpen en opinies in sociale media berichten te

analyseren. De mogelijkheden van digitale tekstanalyse gaan echter nog verder dan het louter testen van bestaande theorieën. Onze ambitie bestaat erin om de nieuwe data mogelijkheden ook te gebruiken om aan theorie ontwikkeling te doen door op een inductieve manier naar onderliggende patronen te zoeken. Door gebruik te maken van data mining technieken op de digitale gebruiker gegevens kunnen we tot inzichten komen over welke gebeurtenissen, personen en onderwerpen burgers 'liken' en willen 'delen'. Concreet, zullen we één belangrijk gepland event bestuderen, de Belgische verkiezingscampagne van 2019, en één niet gepland of onverwacht event in de loop van de volgende jaren. We verwachten dat de informatiestroom voor beide types van events structureel verschillen. Voor elk event plannen we een survey en een grote kwantitatieve dataverzameling gedurende vier weken van alle politieke nieuwsberichten van de belangrijkste nieuwswebsites, en sociale mediaplatformen Twitter en Facebook.

Output: verwacht: sentimentanalyse software voor politieke tekstanalyse, software voor analyse en detectie van “framing” en controversen.

93. Titel: Neurale netwerken en de akoestische identificeerbaarheid van kinderen met een cochleair implantaat. (2017-2021)

Financiering: BOF-UA

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (ACDC)

Beschrijving: Ongeveer 1 op 1000 pasgeborenen wordt gediagnosticeerd met bilateraal, ernstig tot zeer ernstig gehoorverlies. Gehoorapparaten, zoals cochleaire implantaten (CI), hebben ongekende perspectieven gecreëerd voor deze kinderen. Hoewel CI in het algemeen leiden tot een opmerkelijke vooruitgang in de gesproken taalvaardigheid van slechthorende kinderen, blijft hun spraak niettemin vaak afwijken van die van normaal horende kinderen, zelfs als zij het apparaat al verschillende jaren gebruiken. Volwassen sprekers zijn dan ook in staat om het onderscheid te maken tussen de spraak van deze CI-kinderen en die van normaal horende kinderen. Verrassend genoeg zijn de precieze spraakeigenschappen waarop volwassenen zich baseren om tot dergelijke oordelen te komen tot dusver ongrijpbaar gebleven: dat maakt het moeilijk voor klinici om specifieke behandelingen te ontwikkelen. In dit project willen we recente ontwikkelingen in het computationeel leren van "diepe representaties" gebruiken om vat te krijgen op deze kenmerken. Connectionistische modellen (zoals neurale netwerken) hebben veelbelovende resultaten laten zien in het modelleren van ruwe audiosignalen, zoals opgenomen spraak. Door de zorgvuldige analyse, visualisatie en interpretatie van dergelijke modellen, willen we achterhalen welke specifieke eigenschappen in de spraak van CI-kinderen verantwoordelijk zijn voor de identificeerbaarheid van hun spraaksignaal.

Output: software voor gesproken taalverwerking

94. Titel: De maat van het Middelnederlands: reconstructie van het ritme en de prosodie van Middelnederlandse literatuur, een datagedreven benadering. (2015-2019)

Financiering: FWO (aspirantmandaat)

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (ACDC)

Beschrijving: Wat bedoelen we wanneer we zeggen dat het ritme van een literaire tekst "hortend" of "vloeiend" is? En wat zijn de eigenschappen van letterkundige werken die "makkelijk in het geheugen blijven hangen"? De ritmische eigenschappen van literatuur worden doorgaans met intuïtieve en vage termen beschreven en dat is zeker het geval voor de Middelnederlandse letterkunde. De vele berijmde teksten uit de vroegste periode van onze literatuurgeschiedenis worden vaak beschreven in termen zoals de hierboven genoemde. Maar vaak is het onduidelijk wat er precies mee wordt bedoeld. Met dit onderzoek willen we deze intuïtieve en daardoor ook mogelijk subjectieve beweringen wetenschappelijk toetsen en expliciteren. Voor het eerst zullen we computationele technieken aanwenden om de ritmische eigenschappen van Middelnederlandse teksten te onderzoeken. Aangezien deze technieken onbevooroordeeld zijn, kan subjectiviteit worden uitgeschakeld. Zo moet het mogelijk zijn een precies en expliciteerbaar inzicht te krijgen in de ritmische eigenschappen van literaire teksten. Voor de eerste keer zullen we de redenen voor bepaalde intuïtieve observaties precies kunnen beschrijven. We zullen ons daarbij niet beperken tot de analyse van individuele teksten, maar ook de ritmes in diverse literaire genres met elkaar vergelijken. Zonder te verdwalen in een jungle van vage impressies zullen we bv. in staat zijn aan te geven op welke wijze beroemde teksten als "Vanden vos Reynaerde" en "Karel ende Elegast" zich in ritmisch opzicht van elkaar onderscheiden.

Output: software in verband met prosodie

95. Titel: InterStylar: Een stylometrische benadering van intertextualiteit in 12e-eeuwse Latijnse literatuur. (2016-2020)

Financiering: UA-BOF

Partners:

- a. UAntwerpen (CLiPS)
- b. UAntwerpen (ACDC)

Beschrijving: In studies naar auteurschap gebruiken geleerden kwantitatieve technieken uit de stylometrie om anonieme teksten aan gekende auteurs toe te schrijven. Intertextualiteit - het fenomeen waarbij auteurs het werk van andere auteurs in hun eigen werk integreren of erop alluderen -- stelt een interessant probleem in dit verband: moeten alle 'interteksten', zoals citaties of allusies, uit een tekst verwijderd worden voor we de stijl ervan betrouwbaar kunnen analyseren? Dit project gaat in tegen de gangbare opvatting dat dergelijke 'Fremdkörper' enkel ruis zijn en wil de hypothese verifiëren dat intertextualiteit een cruciaal aspect vormt van de individuele schrijfstijl van auteurs. Hiervoor zullen we een representatief corpus analyseren van 12e-eeuwse Latijnse literatuur rondom het indrukwekkende oeuvre van Bernardus van Clairvaux, waarin bijbelse (en andere) intertextualiteit een dominante rol speelt. Het project zal gebruik van de recente vooruitgang in het gebied van 'Deep Representation Learning' oftewel technieken die toelaten om tekstpassages te modelleren als een rauwe reeks letters.

Output: software voor auteursherkenning en intertextualiteitsdetectie

96. Titel: STARLING: Statistisch relationeel leren van natuurlijke taal (2010-2013). Zie [hierboven](#)

C. UNIVERSITEIT GENT – OVERZICHT PROJECTEN

97. Titel: SAPIM - Smart audio processing for interactive multimedia (2010-2011)

Financiering: IWT

Partners: UGent (ELIS-DSL)

Beschrijving: Recent scientific and technological developments have given rise to algorithms that understand to some extent what is happening in a musical audio. In this project we investigate how to adapt them to make them useable in realistic innovative applications. In particular, the project deals with sound recognition (e.g. needed to develop games involving music as an active component), real-time tempo tracking (e.g. to adapt computer generated accompaniment to the tempo of the user ad not vice versa), concatenative music synthesis, etc.

98. Titel: TTNWW: TST Tools voor het Nederlands als Webservices in een Workflow (2010-2012).
Zie hierboven

99. Titel: ORGANIC - Self-organized recurrent neural learning for language processing (2009-2012)

Financiering: EU-FP7

Partners:

- a. UGent (ELIS-DSL)
- b. JU Bremen
- c. Inserm
- d. Sciences U-Lyon
- e. Planet GmbH

Beschrijving: ORGANIC adopts principles of cortical architectures and self-organizing neuro-dynamics for the design of a new type of cognitive architectures. The principle innovations are deep multilevel learning based on recurrent networks using purely dynamic representations and trained both supervised and unsupervised. These techniques will be applied to large vocabulary speech recognition in noisy environments, end on handwriting recognition.

100. Titel: RECAP - Reservoir Computing for auditory pattern recognition (2009-2012)

Financiering: FWO

Partners:

- a. UGent (ELIS-DSL)
- b. UAntwerpen (H. Peremans)

Beschrijving: The aim of this project is to carry out fundamental research into the possibilities and difficulties of applying RC to some challenging auditory pattern recognition applications such as spatial environment recognition from a multi-channel active or passive sonar signal and continuous speech recognition in background noise.

101. Titel: STON: Spraak- en Taaltechnologisch Ondertitelen in het Nederlands (2014-2016). Zie [hierboven](#)

102. Titel: Global TelehomeCare (2015-2017)

Financiering: IWT (KMO innovatie): € 50,000.00 (deelbudget UGent)

Partners:

- a. UGent (ELIS-DSL)
- b. Sebeco
- c. KU Leuven (LUCAS)
- d. HoGent
- e. VUB

Beschrijving: Global Telehomecare is an IWT innovation project that aims at installing and evaluating a novel callcenter-based health monitoring service. Daily telephone conversations with a nursing assistant will counteract social isolation of the (elderly) persons that rely on in-home care. At the same time, the conversations will be used to actively monitor the well-being of the elderly persons.

103. Titel: ASIMOV - Adaptive Semantics Induction for Multilingual matching in an enriched human resOurces enVironment (2016-2018)

Financiering: iMinds ICON

Partners:

- a. UGent (ELIS-DSL) - coördinator
- b. Crosslang
- c. Actonomy
- d. De Persgroep
- e. KU Leuven (ITEC)

Beschrijving: ASIMOV wil betere componenttechnologieën ontwikkelen voor semantische job-matchingsystemen, zodat de juiste kandidaten automatisch worden gelinkt met bepaalde vacatures. Om dat te bereiken, combineert ASIMOV niet alleen Big Data-analyses met natuurlijke taalverwerkingsprocessen en machinaal leren om zo automatisch nieuwe ontologieën af te leiden uit meertalige tekstuele gegevens. Het project verkent ook de semantische interpretatie van andere profielgerelateerde gegevens, zoals online zoekpatronen.

104. Titel: COFI: Comprehensive Integration of Linguistic Features in a Human Inspired Speech Recognition Architecture (2017-2020)

Financiering: BOF

Partners:

- a. UGent (ELIS-DSL)
- b. UGent (LT3)

Beschrijving: In this project, a novel speech recognition framework and a set of accompanying, contemporary linguistic models is co-developed. The envisioned layered architecture combines local inference with inter-layer message passing, similar to what is believed to underlay human speech recognition. By facilitating the incorporation of knowledge (linguistic or other), such architecture is better equipped to model the complex inter-dependencies in speech.

105. Titel: SubTle: SUBjectivity Tagging and Learning (2010-2014)

Financier: Hogeschoolfonds (Hogeschool Gent)

Partners:

- a. UGent (LT3)
- b. Centrum ter Preventie van Zelfdoding

Beschrijving: In het SubTLe-project werd onderzoek verricht naar de automatische detectie van suïcidaliteit op sociale media. Het project resulteerde in een annotatieschema en een

dataset die werd opgebouwd in nauwe samenwerking met het CPZ. Op basis van die data werd een classificatiesysteem gebouwd voor 1) automatisch suïcidaliteitsdetectie en 2) het bepalen van de ernst van een suïciderisico. Met vervolfinanciering van ING en Rotary Foundation werd ook een API gebouwd.

Output:

- annotatieschema
- geannoteerde data
- software
- API.

Outreach:

- Lezingen in De Krook, voor serviceclubs en andere verenigingen over het belang van suïcidaliteitsdetectie op sociale media (vaak samen met de directrice van het Centrum ter Preventie van Zelfdoding
- Lezingen hebben geleid tot bijkomende financiering van de ING bank en van Rotary International voor de ontwikkeling van een API.

106. Titel: Stylene, een webservice voor stylometrie- en leesbaarheidsonderzoek voor het Nederlands (2010-2012). Zie [hierboven](#)

107. Titel: Koppeling van schrijfprocesdata aan lexica (2010-2012)

Financier: FWO

Partners:

- a. UAntwerpen (Luuk Van Waes)
- b. UGent (LT3)

Beschrijving: De tool Inputlog die werd ontwikkeld aan de Universiteit Antwerpen logt het schrijfproces van een bepaalde schrijver (vb. pauzes, inserties, etc.). In dit project werd PoS-tagging en lemmatizing toegevoegd aan Inputlog met het oog op een beter begrip van de lexicale en grammaticale kenmerken van de geschreven tekst.

Output: software

108. Titel: HENDI, een Automatische Helderheidstool voor Engels en Nederlands Discours in Parallele en Vergelijkbare Teksten (2009-2012)

Financier: Hogeschoolfonds

Partners: UGent (LT3)

Beschrijving: In het HENDI-project werd onderzoek verricht naar een nieuwe holistische leesbaarheidsmetriek die gebruik maakt van lerende technieken als alternatief voor de verouderde leesbaarheidsformules die enkel rekening houden met oppervlakkige tekstkarakteristieken. Voor het verzamelen van de data werd niet alleen gebruik gemaakt van experts, maar ook van crowdannotaties.

Output:

- geannoteerde corpora
- software en demo's (zie LT3-website).

Outreach: publieke demo i.v.m. leesbaarheid

109. Titel: ParaSense or how to use parallel corpora for Word Sense Disambiguation (2007-2012)

Financier: Hogeschoolfonds

Partners: UGent (LT3)

Beschrijving: In dit project werd onderzoek verricht naar automatische desambiguering van woordbetekenissen op basis van parallelle corpora.

Output:

- data voor een SemEval shared task
- software en demo's.

110. Titel: TExSIS, Terminology Extraction for Semantic Interoperability and Standardization (2011- 2012)

Financier: IWT Tetra

Partners:

- a. LT3 (UGent)
- b. IBCN (UGent)

c. Een gebruikerscommissie van bedrijven

Beschrijving: In het TEXSIS-project werd een nieuw systeem ontwikkeld voor automatische terminologie-extractie op basis van monolinguale en bilinguale data. Het gaat hier om een hybride termextractor die zowel gebruik maakt van statistische filtering als van automatische linguïstische preprocessing.

Output:

- geannoteerde corpora
- software

111. Titel: PARIS, Personalised AdveRtisements built from web Sources (2012-2016). Zie boven

112. Titel: SentiFM, sentiment mining for the financial markets (2012-2015)

Financier: IWT (SB)

Partners: LT3 (UGent)

Beschrijving: In het SentiFM-mandaatsonderzoek werd onderzoek verricht naar de extractie van economische events in financiële literatuur en het daaraan gekoppelde (impliciet) sentiment.

Output: software

113. Titel: CHLG. A parsed corpus of Historical Low German (2014-2018)

Financier: Hercules

Partners:

- a. LT3 (UGent)
- b. Anne Breitbarth (UGent)

Beschrijving: In dit project werd een corpus samengesteld van Middelnederduits. Naast de verzameling van de data lag de focus vooral op de (deels automatische) syntactische annotatie van de data. Voor dit project werd een nieuwe PoS tagger ontwikkeld en een parser.

Output: software

114. Titel: SCATE: Smart Computer-Aided Translation Environment (2014-2018). Zie boven

115. Titel: Comprehensive Integration of Linguistic Features in a Human Inspired Speech Recognition Architecture (2017-2021)

Financier: BOF UGENT (interdisciplinair onderzoeksproject)

Partners:

- a. LT3 (UGent)
- b. ELIS (UGent)

Beschrijving: Dit project wil via taalmodellering op verschillende niveaus (morfologisch, lexicaal, syntactisch, semantisch) input leveren aan een neurale ASR-systeem.

Output: software

116. Titel: EXTRACT. Extraction of terminology from comparable texts (2017-2021)

Financier: FWO

Partners: LT3 (UGent)

Beschrijving: Vertalers maken voor hun vertalingen gebruik van bilinguale termenbanken die ze nu vaak nog manueel aanleggen. Automatische terminologie-extractie (ATE) kan hen hierbij ondersteunen, maar dat gebeurt op dit ogenblik vooral nog op basis van schaarse parallelle corpora. In dit project gaan we na hoe ATE kan geleerd worden op basis van niet-parallelle, vergelijkbare corpora.

Output: software en demo's.

117. Titel: AMiCA: Automatic Monitoring for Cyberspace Applications (2013-2016). Zie boven

118. Titel: SENTiVENT: event extraction and sentiment analysis for financial news (2017-2020)

Financier: FWO

Partners: LT3 (UGent)

Beschrijving: In het SENTiVENT-project wordt onderzocht hoe economische events kunnen gemodelleerd en geëxtraheerd worden (granulariteit, classificatie, etc.). Er wordt ook nagegaan in hoeverre automatische eventextractie fluctuaties in beurskoersen kan voorspellen.

Output: software en demo's.

119. Titel: INSIGHT: Intelligent Neural Systems as InteGrated Heritage Tools (2016-2021). Zie ***Fout! Verwijzingsbron niet gevonden.***

120. Titel: TAPAS (Training Network on Automatic Processing of PAtHological Speech) (2017-2020)

Financier: EU Horizon 2020 (MSCA - ITN – ETN)

Partners:

- a. UGent (imec-IDLab)
- b. Idiap Research Institute
- c. Friedrich-Alexander-Universität Erlangen Nuernberg
- d. INESC ID - Instituto de Engenharia de Sistemas E Computadores Investigacao E Desenvolvimento em Lisboa
- e. Ludwig-Maximilians-Universität Muenchen
- f. The Netherlands Cancer Institute - Antoni van Leeuwenhoek
- g. Philips Research Eindhoven
- h. Radboud Universiteit Nijmegen
- i. Universität Augsburg
- j. Université Toulouse III-Paul Sabatier.

Beschrijving: The TAPAS work programme targets three key research problems:

(a) Detection: developing speech processing techniques for early detection of conditions that impact on speech production. The outcomes will be cheap and non-invasive diagnostic tools that provide early warning of the onset of progressive conditions such as Alzheimer's and Parkinson's.

(b) Therapy: Using newly-emerging speech processing techniques to produce automated speech therapy tools. These tools will make therapy more accessible and more individually

targeted. Better therapy can increase the chances of recovering intelligible speech after traumatic events such a stroke or oral surgery.

(c) Assisted Living: Re-designing current speech technology so that it works well for people with speech impairments and also helps in making informed clinical choices. People with speech impairments often have other co-occurring conditions making them reliant on carers. Speech-driven tools for assisted-living are a way to allow such people to live more independently.

D. VRIJE UNIVERSITEIT BRUSSEL – OVERZICHT PROJECTEN

121. Titel: Language games voor het taalonderwijs (2009-2013)

Financiering: IWT – SBO

Partners: Artificial Intelligence Lab, Vrije Universiteit Brussel

Beschrijving: The project aimed at offering language learners a platform to optimize their knowledge of particular language system (e.g. Russian aspect, Spanish modality, etc.) through interactions with artificial agents in order to reach optimal communicative success and to improve the general cognitive capacities of the language learner. The existing language game framework that is typically used for agent - agent interactions, is been extended here to host human - agent communication.

Output:

- FCG grammar for the Spanish verb phrase
- Babel2 language game framework extended to work for human-computer interaction
- doctoraat Katrien Beuls

122. Titel: Digging for the roots of understanding (DRUST) (2011-2014)

Financiering: European Science Foundation

Partners:

- a. Professor Luc Steels (VUB)
- b. Dr. Andreas Roepstorff (University of Aarhus)
- c. Professor Dan Zahavi (University of Copenhagen)
- d. Professor Claudio Castellano (Universita di Roma)
- e. Dr. Joanna Raczaszek-Leonardi (University of Warsaw).

Beschrijving: The Collaborative Research Project DRUST brings together eminent European research groups that cover the full breadth of Cognitive Science (including cognitive anthropology, cognitive neuroscience, cognitive psychiatry, artificial intelligence, linguistics, philosophy, and psychology) in an interdisciplinary examination of how different types of common ground support interpersonal and intercultural understanding. Communication plays a key role for such an understanding but only to the extent that dialog partners have sufficient common ground. Misunderstandings occur because most dimensions on which common ground can exist are not universally shared and not a priori present. New research in the Humanities and the Cognitive Sciences has shown that partners in dialogue and partners in action often differ in their perception of the situation, prior background, inference

capacity, and linguistic conventions. Although common ground is available when individuals interact with members of their own culture in well-defined contexts, considerable gaps in common ground occur in today's multicultural societies where individuals with different cultural backgrounds interact in heterogeneous contexts. Profound misunderstandings easily go unnoticed. A better comprehension of the neural, psychological, and cultural processes that establish common ground is thus crucial for identifying the roots of interpersonal understanding and intercultural understanding and can help to reduce misunderstandings and the serious consequences such misunderstandings can cause.

123. Title: Evolution of Shared Semantics in Computational Environments (ESSENCE) (2013-2017)

Financiering: Marie Curie Initial Training Network

Partners:

- a. Artificial Intelligence Lab, Vrije Universiteit Brussel
- b. School of Informatics, The University of Edinburgh
- c. Institute for Logic, Language, and Computation, Universiteit van Amsterdam
- d. Department of Information Engineering and Computer Science, Universita di Trento
- e. SONY Computer Science Labs Paris; TAIGER Madrid

Beschrijving: ESSENCE was a European Marie Curie Actions Project (FP7, ITN) that aims to improve robustness and resilience of natural language processing systems. In everyday life, humans exhibit strong skills in resolving communication problems by re-negotiating what they mean. Modern-day computational systems, however, are lacking in resilience and robustness in this respect. Whenever different components with different vocabularies and models of meaning interact within distributed, open environments, they have to rely on their human designers' abilities to resolve problems of miscommunication. The aim of the ESSENCE network is to translate these abilities of natural communicating systems to computational systems in order to improve their resilience and robustness, and by doing so to also make these systems more comprehensible to human users. To this end, we will bring together research on phenomena observed in human communication with research from areas that investigate heterogeneous computational communicating systems.

Output: huge improvements to the FCG formalism and engine that implements constructional processing of natural language sentences (see code repository on GitHub: <https://github.com/EvolutionaryLinguisticsAssociation/Babel2> and new notation in Steels, Luc and Katrien Beuls (Eds.) (2017) Fluid Construction Grammar and the Verb Phrase. Special issue in Constructions and Frames 9(2)).

124. Titel: Artificial Language Understanding In robotS (ATLANTIS) (2016 – 2018)

Financiering: CHIST-ERA (Human Language Understanding call 2015)

Partners:

- a. Artificial Intelligence Lab, Vrije Universiteit Brussel
- b. Austrian Research Institute for AI (OFAI)
- c. Lattice – CNRS
- d. Universitat Pompeu Fabra
- e. Sony Computer Science Labs

Beschrijving: ATLANTIS attempts to understand and model the very first stages in grounded language learning, as we see in children until the age of three: how pointing or other symbolic gestures emerge from the ontogenetic ritualization of instrumental actions, how words are learned very fast in contextualized language games, and how the first grammatical constructions emerge from concrete sentences. This requires a global, computational theory of symbolic development that informs us about what forces motivate language development, what strategies are exploited in learner and caregiver interactions to come up with more complex compositional meanings, how new grammatical structures and novel interaction patterns are formed, and how the multitude of developmental pathways observed in humans lead to a full system of multi-modal communication skills. This ambitious aim is feasible because there have been very significant advances in humanoid robotics and in the development of sensory-motor competence recently, and the time is ripe to push all this to a higher level of symbolic intelligence, going beyond simple sensory-motor loops or pattern-based intelligence towards grounded semantics, and incremental, long-term, autonomous language learning.

Output:

- new robot interface for Nao (Common Lisp)
- VR data set recorded about NL instructions people give to navigate in the virtual world aligned with the world models that we get from the Unity software

125. Titel: ODYCCEUS (2017 – 2020)

Financiering: H2020 Fet Proactive

Partners:

- a. Vrije Universiteit Brussel, Belgium (Artificial Intelligence Lab)
- b. Max Planck Institute for Mathematics in the Sciences, Germany
- c. Università Ca'Foscari, Venice, Italy (Laboratory for Experimental Economics & Department of Management and Department of Humanities)

- d. Chalmers University of Technology, Gothenburg, Sweden (Complex Systems Group at the division of Physical Resource Theory, Department of Energy and Environment)
- e. Université Pierre et Marie Curie - Paris 6, France (Computer Science Laboratory)
- f. Université Paris Diderot - Paris 7, France (Department of Geography History and Economy - GIS CIST & UMR Géographie)
- g. Universiteit van Amsterdam, Netherlands (Department of Sociology at Faculty of Social and Behavioural Sciences and Digital Methods Initiative)
- h. Universität Leipzig, Institut für Soziologie, Germany

Beschrijving: The project aims at developing new empirically grounded theoretical foundations and new interactive and participatory tools for understanding, analyzing, visualizing and monitoring the emergence of social conflicts and crises driven by differences in culture and by diverging world-views, with a focus on their distribution in the European geopolitical and media space. It will also develop methods and technologies for the early detection and localization of potential conflict areas through the analysis of digital media and online communities, and the participatory development of new arenas of discourse supporting conflict resolution in collaboration with a group of stakeholders. On the theoretical ground the project aims at developing new concepts, mathematical foundations and computational tools for understanding and modeling social conflict and crises in systems that have high dimensionality, little decomposability, and generate non stationary processes. As such, it promises to deliver new insights on the “global systems science” of conflict. Key theoretical contributions will include new geometric models of partial representations, with applications to the representation of political spaces and game theory models of cultural conflict. The project will also deliver new computational models for the analysis of natural language that may allow a more precise and theory-driven analysis. Finally, the project will develop new models of the spatiotemporal distribution of conflict and crises.

The theoretical foundations will provide the basis for the development of innovative interactive (participatory) tools to monitor, visualize and model the dynamics of conflict situations that involve heterogeneous cultural biases and non-transparent entanglements of multilateral interests. We aim to use diverse empirical sources (from digitalized historical corpora on 19th and 20th centuries to real-time online communication) to understand the mechanisms that shape the discourse around controversial issues and which hinder or foster a process of compromise. The use cases studied in the project will include migration and the refugee crises, new populist movements in Europe, climate change debates, tensions in external relations (with neighbors) and internal relations (separatist movements) of the European Union, and the emergence of racist rhetoric. On the other hand, the internet also provides an ecology in which new forms of interaction may be established at pace. We aim at enriching this ecology with an online tool which maps ongoing discourses onto empirically informed representations that allow users (and stakeholders) to grasp the complexity of the problem in question and, in addition to that, provides the possibility of negotiating the importance and valence of key issues in a participatory fashion. We believe that an empirically informed and collectively negotiated framing of a debate can be a basis for new forms of participatory social decision making in which the media sphere becomes a true democratic element.

NETWERKEN, VALORISATIE EN MAATSCHAPPELIJKE BENUTTING

In het kader van de vele projecten die in het vorige hoofdstuk zijn gepresenteerd, werden niet alleen vele publicaties (zie *Appendix*: overzicht publicaties) en resources (data en tools) geproduceerd, maar werden ook strategische nationale en internationale netwerken opgezet met kennisinstellingen, bedrijven, overheidsorganisaties en maatschappelijke partijen, met groot potentieel voor valorisatie en maatschappelijke benutting van de resultaten. Interessante voorbeelden hiervan zijn de netwerken die KU Leuven via COST-acties heeft opgezet rond verschillende thema's. **iv&L Net** (Europees netwerk voor de integratie van visie en taal, EU ICT COST Action IC1307, KU Leuven en 28 Europese landen) brengt onderzoekers uit de academische wereld en de industrie samen die beeld en taal verwerken. Deze nieuwe discipline draagt bij tot het automatisch interpreteren van beeld en van natuurlijke taal, alsook verschaft het nieuwe inzichten in de cognitie en universele representaties van kennis. Er zijn daarnaast talrijke applicaties zoals multimodaal en cross-modaal zoeken van multimedia. De output bestaat uit een archief van publicaties, software en data: <http://ivl-net.eu/repositories/>. **TextLink** (Het structureren van het discours in een multilinguaal Europa, EU ISCH COST Action IS1312, KU Leuven met 26 Europese landen en 3 intercontinentale partners), promoot standaardannotaties en tools voor discoursstudies in Europese talen. Output bestaat uit een archief van publicaties, tools en data: <http://textlink.ii.metu.edu.tr/resources-0>. **MUMIA** (Multilinguale en multigefacetteerde interactieve informatietoegang), EU ICT COST Action IC1002, KU Leuven met 28 Europese landen en 4 intercontinentale partners) heeft betrekking op zoektechnologie die in staat moet zijn om meerdere en gedistribueerde bronnen van meertalige en veelzijdige inhoud te zoeken, filteren, classificeren, extraheren, combineren, integreren en verwerken. Ook in dit geval bestaat de output uit een archief van publicaties, tools en data: <http://www.mumia-network.eu/index.php/resources/dissemination> en <http://www.mumia-network.eu/index.php/resources/datasets>.

Via het Europees project "Able to Include" (EU, CIP PSP) heeft KU Leuven contacten gelegd met belangrijke Europese spelers (The National Microelectronics Applications Centre (MAC), Ierland, Thomas More Kempen, Universitat Pompeu Fabra, Building Bridges Training, Inclusion Europe, en Fundación Prodis) met als doel de leefomstandigheden van mensen met een intellectuele beperking te verbeteren, onder meer door hen ook toegang te geven tot de social media (e-inclusion) en de toepassingen die ontwikkeld zijn binnen dit netwerk worden nu daadwerkelijk gebruikt door bv. organisaties als WAI-NOT.

In het kader van het project PARIS (Personalised Advertisements built from web Sources, (Gepersonaliseerde advertenties gebouwd uit webbronnen) is sentimentanalyse software ontwikkeld dat vervolgens werd gecommercialiseerd in een VLAIO-innovatieproject voor het bedrijf Hello Customer. Onderdelen van SCATE (terminologie-extractie) zijn ook verder ingezet in dienstverleningsprojecten voor EBM Practicenet, en voor TheMatchBox (in het kader van een VLAIO haalbaarheidsstudie). Van de verschillende projecten werden demo's ontwikkeld die nu op de website van LT3 staan (leesbaarheidsdemo's: Hendi, sentimentanalyse: PARIS).

UAntwerpen is coördinator van WOG DHBelux en werkt samen met verschillende start-ups en gevestigde bedrijven op het gebied van computertaalkunde en Digital Humanities.

In het kader van het AMiCA-project heeft UGent een strategisch netwerk opgezet met serviceclubs en andere verenigingen voor suïcidaliteitspreventie. Dit heeft ook geleid tot bijkomende financiering van de ING bank en van Rotary International voor de ontwikkeling van een API. Naar aanleiding hiervan zijn verschillende lezingen gehouden samen met de directrice van het Centrum ter Preventie van Zelfdoding, is een perscampagne gelanceerd waarbij opgeroepen werd om cyberpestdata te delen, en er zijn interviews gegeven voor het nieuws op Radio2, het topic was ook een item in het VTM-nieuws en uiteindelijk is er afsluitende workshop georganiseerd met vele bedrijven, relevante organisaties en een panel gemodereerd door Jan Hautekiet.

TOEKOMSTIGE ONTWIKKELINGEN: CLARIAH-VL

Het staat vast dat CLARIN-ERIC een bijzonder nuttig netwerk is voor onze Vlaamse onderzoekers. Het bovenstaande rapport toont aan dat de onderzoeksgroepen van het consortium zeer actief zijn binnen de taal- en spraaktechnologie en zeer succesvol zijn bij projectonderzoek en valorisatie. Het laat de Vlaamse onderzoekers ook toe om deel te nemen aan het internationale netwerk van taal- en spraaktechnologie.

Nu is het tijd om een stap verder te gaan: de verworvenheden van de taal- en spraaktechnologie en de expertise van de computerwetenschappers, de computerlinguïsten en andere gerelateerde disciplines wordt nu verder uitgebreid naar het volledige forum van de “arts and humanities”. Digital humanities is nu aan zet: overall ziet men de resultaten van het netwerk DARIAH-EU prominent naar voren komen. Vlaanderen kan hier niet achter blijven.

Kijken we naar Nederland, dan zien we dat CLARIAH³ (Common Lab Research Infrastructure for the arts and the humanities) tot mooie resultaten leidt. Ik citeer hier enkele feiten over het Nederlandse consortium CLARIAH:

“Bij de vervolgaanvraag van CLARIN-NL (2009-2015) werd besloten de krachten te bundelen tot een project: CLARIAH, dit als een deel van DARIAH. De CLARIAH bijdrage in DARIAH wordt besteed aan activiteiten om de missie in Nederland te versterken en in staat zijn om actief deel te nemen aan de Europese DARIAH activiteiten.

CLARIAH Nederland startte met een voorbereidend project binnen de Nationale Roadmap Grootschalige onderzoeksinfrastructuur. Dit leidde tot een grotere aanvraag (2014-18) waarin sterker kon worden gewerkt op infrastructuur, taalkunde en andere realisaties;

Er is een nauwe samenwerking, zowel op nationaal als op internationaal niveau tussen de CLARIN-EU en DARIAH-EU-projecten, zoals blijkt uit de gezamenlijk georganiseerde conferenties zoals NEERI en de ondersteuning van Digital Humanities. Het is dan ook logisch om de krachten verder te bundelen en met elkaar in een gemeenschappelijk project te versterken, zoals CLARIAH.

CLARIAH is verder betrokken bij een aantal internationale netwerken die rechtstreeks ten dienste staan van digital humanities communities, zoals DASISH, EUDAT, en de Research Data Alliance. Bovendien zijn verscheidene CLARIAH-partners betrokken geweest bij het opzetten van de uitwisseling faciliteiten zoals META-SHARE in de Europese META-NET-project en de Europese Holocaust Research Infrastructure (EHRI). Daarnaast spelen verscheidene CLARIAH-partners een leidende rol in Europeana en aanverwante "Digitaal Erfgoed" projecten, die een meertalige online verzameling van miljoenen gedigitaliseerde items uit Europese musea, bibliotheken, archieven en multimedia-collecties bieden.”⁴

Ook voor Vlaanderen is de vraag vanuit de geesteswetenschappen groot: historici, letterkundigen, kunsthistorici, specialisten mediastudies en socio-economische geschiedenis willen toegang tot belangrijke digitale collecties en willen de beschikbare infrastructuur kunnen gebruiken voor nieuw

³ CLARIAH, als een nationale tegenhanger van CLARIN-EU en DARIAH-EU, is direct en diep verankerd in de Europa-brede ESFRI-onderneming. CLARIN en DARIAH zijn de enige twee, op de geesteswetenschappen gerichte infrastructuurprogramma's op de ESFRI-Roadmap.

⁴ <https://www.clariah.nl>

onderzoek. De enorme hoeveelheid beschikbare digitale bronnen (tekst, audiovisueel en gestructureerde data) betekent een revolutie in de geesteswetenschappen. Toponderzoek is alleen mogelijk mits gebruik te maken van gesofisticeerde ICT-hulpmiddelen. CLARIAH kan daarbij een gemeenschappelijke onderzoeksinfrastructuur leveren die het mogelijk maakt om grote collecties van digitale data te verwerken en innovatieve gebruiksvriendelijke software te gebruiken voor de analyse van deze data.

Zowel de data als de tools zullen duurzaam worden beheerd, zodat ze ook in de toekomst beschikbaar zijn voor alle geïnteresseerde gebruikers, dit uiteraard via een CLARIN-centrum. Van taalkundigen tot mediawetenschappers en van historici tot spraaktechnologen, alle onderzoekers in de humanities zijn betrokken maar het consortium moet ook de brug leggen naar de bedrijfswereld, musea, bibliotheken en publieke diensten.

APPENDIX: OVERZICHT PUBLICATIES

Spyns, P. and Odijk, J. (Editors), *Essential Speech and Language Technology for Dutch: Results by the STEVIN programme*. Springer, 2013, xvii + 413 pp.

Odijk, J. and van Hessen, A. (eds.) *CLARIN in the Low Countries*. London: Ubiquity Press, 2017. DOI: <https://doi.org/10.5334/bbi>. License: CC-BY 4.0

A. KU LEUVEN – OVERZICHT PUBLICATIES

Alfarone, D., Davis, J., "Unsupervised learning of an IS-A taxonomy from a limited domain-specific corpus", in *_Proceedings of the 24th International Joint Conference on Artificial Intelligence_ (25-31 July 2015, Buenos Aires, Argentina), 2015, pp. 1434-1441.*

Annema, J., Poels, Y., Zaman, B., Cornillie, F., "Design of Web-Based Mini-Games for Language Learning: An Evidence-Based and User-Centred Approach" in *Serious Games: The Challenge*, De Wannemacker, S., Vandercruysse, S. Eds., Berlin, Germany: Springer Verlag, 2012, ISBN 978-3-642-33813-7, pp. 79-81.

Antoniadis, G., Desmet, P., "NLP for learning and teaching: challenges and opportunities", *Traitement Automatique des Langues*, vol. 57, no. 3, 2016, pp. 7-13.

Asoodar, M., Marandi, S., Vaezi, S., Desmet, P., "Podcasting in a virtual English for academic purposes course: learner motivation", *Interactive Learning Environments*, vol. 24, no. 4, 2016, art.nr. 10.180/10494820.2014.937344, pp. 875-896.

Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., Bernal-Delgado, E., Blomberg, N., Bock, C., Conesa, A., Del Signore, S., Delogne, C., Devilee, P., Di Meglio, A., Eijkemans, M., Flicek, P., Graf, N., Grimm, V., Guchelaar, H., Guo, Y., Gut, I., Hanbury, A., Hanif, S., Hilgers, R., Honrado, Á., Hose, D., Houwing-Duistermaat, J., Hubbard, T., Janacek, S., Karanikas, H., Kievits, T., Kohler, M., Kremer, A., Lanfear, J., Lengauer, T., Maes, E., Meert, T., Müller, W., Nickel, D., Oledzki, P., Pedersen, B., Petkovic, M., Pliakos, K., Rattray, M., I Màs, J., Schneider, R., Sengstag, T., Serra-Picamal, X., Spek, W., Vaas, L., van Batenburg, O., Vandelaer, M., Varnai, P., Villoslada, P., Vizcaíno, J., Wubbe, J., Zanetti, G., "Making sense of big data in health research: Towards an EU action plan", *Genome Medicine*, vol. 8, no. 1, 2016, art.nr. 10.1186/s13073-016-0323-y.

Augustinus, L., "Complement Raising and Cluster Formation in Dutch. A Treebank-supported Investigation.", PhD Thesis (Advisor(s): Van Eynde, F.), (Co-Advisor(s): Smessaert, H.), 2015. ISBN 978-94-6093-195-6.

Augustinus, L., Dirix, P., "Dit Het Op(ge)hou Reën." *The IPP Effect in Afrikaans*, (18 January 2013, Enschede), 2013.

Augustinus, L., Dirix, P., van Niekerk, D., Schuurman, I., Vandeghinste, V., Van Eynde, F., van Huyssteen, G., "AfriBooms: An Online Treebank for Afrikaans", in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016) (23-28 May 2016, Portorož)*, 2016, pp. 677-682.

Augustinus, L., Van Eynde, F., "Dat we ons daar nog mee kunnen bezig houden." *Looking for cluster creepers in Dutch treebanks*, (17 January 2014, Leiden), 2014.

Augustinus, L., Van Eynde, F., "A treebank-based investigation of IPP-triggering verbs in Dutch and German", (11-12 January 2013, Leuven), 2013.

Augustinus, L., Van Eynde, F., "A Treebank-based Investigation of IPP-triggering verbs in Dutch", in *Proceedings of the Eleventh International Workshop on Treebank and Linguistic Theories (TLT11) (30 November - 1 December 2012, Lisbon, Portugal)*, Hendrickx, I., Kübler, S., Simov, K. Eds., 2012, pp. 7-12.

Augustinus, L., Van Eynde, F., "A Treebank-Based Investigation of IPP-Triggers in Dutch", (19-21 September 2012, Leuven), 2012.

Augustinus, L., Van Eynde, F., "A usage-based typology of Dutch and German IPP verbs", *Leuvense Bijdragen: Tijdschrift voor Germaanse Filologie*, vol. 101, 2017, pp. 101-122.

Augustinus, L., Van Eynde, F., "Adposition stranding in Dutch. "Omdat we er dol op zijn!", (5-6 February 2015, Antwerp), 2015.

Augustinus, L., Van Eynde, F., "Looking for Cluster Creepers in Dutch Treebanks. Dat we ons daar nog mee bezig houden", *Computational Linguistics in the Netherlands Journal*, vol. 4, 2014, pp. 149-170.

Augustinus, L., Vandeghinste, V., "Querying Treebanks by Example", (20 January 2012, Tilburg), 2012

Augustinus, L., Vandeghinste, V., Schuurman, I., "Lumberjacking with GrETEL", (11 May 2013, Leuven), 2013.

- Augustinus, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Example-Based Treebank Querying with GrETEL – now also for Spoken Dutch", in Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013) (22-24 May 2013, Oslo), Oepen, S., Hagen, K., Bondi Johannesse, J. Eds., 2013, pp. 423-428.
- Augustinus, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Finding your way through the woods with GrETEL", (14 June 2013, Groningen), 2013.
- Augustinus, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "GrETEL: A Tool for Example-Based Treebank Mining" in CLARIN in the Low Countries, Odijk, J., van Hessen, A. Eds., London: Ubiquity Press, 2017, ISBN 978-1-911529-24-8, pp. 269-280.
- Augustinus, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "No longer lost in the forest..", (5-6 February 2015, Antwerp), 2015.
- Augustinus, L., Vandeghinste, V., Van Eynde, F., "Example-Based Treebank Querying", Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012) (21-27 May 2012, Istanbul), 2012, pp. 3161-3167.
- Augustinus, L., Vandeghinste, V., Van Eynde, F., "Facilitating treebank mining of CGN with GrETEL", (18 January 2013, Enschede), 2013.
- Augustinus, L., Vandeghinste, V., Vanallemeersch, T., "Poly-GrETEL: Cross-Lingual Example-based Querying of Syntactic Constructions", in Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016) (23-28 May 2016, Portorož), 2016, pp. 3549-3554.
- Augustinus, L., Vanroy, B., Vandeghinste, V., "Improved Treebank Querying: A Facelift for GrETEL", in (18-20 September 2017, Budapest), 2017.
- Baby, D., "Non-negative Sparse Representations for Speech Enhancement and Recognition", PhD Thesis (Advisor(s): Van hamme, H.), 2016.
- Baby, D., Gemmeke, J., Virtanen, T., Van hamme, H., "Exemplar-based speech enhancement for deep neural network based automatic speech recognition", in Proceedings ICASSP 2015 (19-24 April 2015, Brisbane, Australia), 2015, pp. 4485-4489.
- Baby, D., Van hamme, H., "Hybrid input spaces for exemplar-based noise robust speech recognition using coupled dictionaries", in Proceedings EUSIPCO 2015 (31 August - 4 September 2015, Nice, France), 2015, pp. 1706-1710.
- Baby, D., Van hamme, H., "Investigating modulation spectrogram features for deep neural network-based automatic speech recognition", in Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany), 2015, pp. 2479-2483.
- Baby, D., Van hamme, H., "Supervised speech dereverberation in noisy environments using exemplar-based sparse representations", in Proceedings ICASSP 2016 (20-25 March 2016, Shanghai, China), 2016, pp. 156-160.
- Baby, D., Virtanen, T., Barker, T., Van hamme, H., "Coupled dictionary training for exemplar-based speech enhancement", in Proceedings ICASSP 2014 (4-9 May 2014, Florence, Italy), 2014, pp. 2907-2911.
- Baby, D., Virtanen, T., Gemmeke, J., Barker, T., Van hamme, H., "Exemplar-based noise robust automatic speech recognition using modulation spectrogram features", in Proceedings SLT 2014 (7-10 December 2014, South Lake Tahoe, Nevada, USA), 2014, pp. 1-6.
- Baby, D., Virtanen, T., Gemmeke, J., Van hamme, H., "Coupled dictionaries for exemplar-based speech enhancement and automatic speech recognition", IEEE-acm Transactions on Audio Speech and Language Processing, vol. 23, no. 11, 2015, pp. 1788-1799.
- Baby, D., Virtanen, T., Van hamme, H., "Coupled dictionary-based speech enhancement with adaptively learned atoms for the CHiME-3 challenge", (9 January 2016, Groningen, The Netherlands), 2016.
- Bahari, M., "Automatic Speaker Characterization; Automatic Identification of Gender, Age, Language and Accent from Speech Signals (Automatische sprekercharacterisatie; Automatische identificatie van geslacht, leeftijd, taal en accent uit stemopnamen)", PhD Thesis (Advisor(s): Van hamme, H., Wambacq, P.), 2014.
- Bahari, M., Dehak, N., Van hamme, H., Burget, L., Ali, A., Glass, J., "Non-Negative Factor Analysis of Gaussian Mixture Model Weight Adaptation for Language and Dialect Recognition", IEEE Transactions on Audio Speech and Language Processing, vol. 22, no. 7, 2014, pp. 1117-1129.
- Bahari, M., McLaren, M., Van hamme, H., Van Leeuwen, D., "Age estimation from telephone speech using i-vectors", in Proceedings Interspeech 2012 (9-13 September 2012 (ISCA Best Student Paper Award), Portland, Oregon, USA), 2012, pp. 506-509.
- Bahari, M., McLaren, M., Van hamme, H., van Leeuwen, D., "Speaker age estimation using i-vectors", Engineering Applications of Artificial Intelligence, vol. 34, 2014, pp. 99-108.

- Bahari, M., Saeidi, R., Van hamme, H., van Leeuwen, D., "Accent recognition using i-vector, gaussian mean supervector and gaussian posterior probability supervector for spontaneous telephone speech", in Proceedings ICASSP'2013 (26-31 May 2013, Vancouver, Canada), 2013, pp. 7344-7348.
- Bahari, M., Van hamme, H., "Normalized ordinal distance; A performance metric for ordinal, probabilistic-ordinal or partial-ordinal classification problems" in Case studies in intelligent computing: achievements and trends, Biju, I., Israr, N. Eds., United States: CRC Press, Taylor and Francis, 2014, ISBN 978-1-4822-0703-3, 978-1-4822-0704-0, pp. 285-302.
- Bahari, M., Van hamme, H., "Normalized ordinal distance; A performance metric for ordinal, probabilistic-ordinal or partial-ordinal classification problems", in Biometric Technologies in Forensic Science (14-15 October 2013, Nijmegen, The Netherlands), 2013, pp. 1-7.
- Bahari, M., Van hamme, H., "Speaker adaptation using maximum likelihood general regression", in (3-5 July 2012, Montreal, Quebec, Canada), 2012.
- Bahari, M., Van hamme, H., "Speaker age estimation using hidden markov model weight supervectors", in (3-5 July 2012, Montreal, Quebec, Canada), 2012.
- Barker, T., Van hamme, H., Virtanen, T., "Modelling primitive streaming of simple tone sequences through factorisation of modulation pattern tensors", in Proceedings Interspeech 2014 (14-18 September 2014, Singapore), 2014, pp. 1371-1375.
- Belz, A., Coheur, L., Ferrari, V., Moens, M., Pastra, K., Vulic, I., Eds, "Proceedings of the workshop on vision and language 2015 (VL'15)" ACL, 2015. ISBN 978-1-941643-32-7.
- Belz, A., Moens, M., Smeaton, A., Eds, "Proceedings of the 3rd annual meeting of the EPSRC network on vision and language and the 1st technical meeting of the European network on integrating vision and language", East Stroudsburg, USA: ACL, 2014. ISBN 978-1-873769-28-1.
- Bertels A., Speelman D. (2016). Analyse sémantique distributionnelle dans un corpus technique : les enjeux sémantiques dans un jeu de statistiques. Actes de JADT 2016. Journées internationales d'analyse statistique des données textuelles (JADT). Nice, 7-10 June 2016 (pp. 273-283).
- Bertels, A., De Hertog, D., Heylen, K., "Etude sémantique des mots-clés et des marqueurs lexicaux stables dans un corpus technique", in Actes de la conférence conjointe JEP-TALN-RECITAL 2012 (4-8 June 2012, Grenoble), 2012, pp. 239-252.
- Bethard, S., Kolomiyets, O., Moens, M., "Annotating story timelines as temporal dependency structures", in Proceedings of the eighth international conference on language resources and evaluation (LREC) (21-27 May 2012, Istanbul, Turkey), 2012, pp. 2721-2726.
- Bibauw, S., François, T., Desmet, P., "Conversational agents for language learning: state of the art and avenues for research on task-based agents", (27-29 May 2015, Boulder, CO), 2015.
- Bibauw, S., François, T., Desmet, P., "Dialogue-based CALL: an overview of existing research", in Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy (26-29 August 2015, Helm, F., Bradley, L., Guarda, M., Thouësny, S. Eds., 2015, pp. 57-64.
- Bibauw, S., François, T., Desmet, P., "Instructional design and natural language processing in dialogue-based CALL", (10-12 May 2016, East Lansing, MI), 2016.
- Bibauw, S., François, T., Desmet, P., "The effectiveness of dialogue-based CALL on L2 proficiency development: a meta-analysis", (16-20 May 2017, Flagstaff, AZ), 2017.
- Bodard, G., H. Cayless, M. Depauw, L. Isaksen, K.F. Lawrence, S. Rahtz, "Standards for Networking Ancient Person-data: Digital Approaches to Problems in Prosopographical Space", in: Digital Classics Online 3.1 (2017), pp. 23-38. Accessible at <http://journals.ub.uni-heidelberg.de/index.php/dco/issue/view/3634>
- Boes, W., Van Rompaey, R., Pelemans, J., Verwimp, L., Wambacq, P., "Domain adaptation for LSTM language models", Book of abstracts CLIN27 (10 February 2017, Leuven, Belgium), 2017, pp. 57.
- Bonachela Capdevila, F., Montero Perez, M., De Wannemacker, S., "Personal dictionary for text prioritization and selection", (17 January 2014, 2014).
- Boyd, K., Santos Costa, V., Davis, J., Page, D., "Unachievable region in precision-recall space and its effect on empirical evaluation", in Proceedings of 29th International Conference on Machine Learning_ (June 26th - July 1st, Edinburgh, Scotland), 2012, pp. 1-9.
- Briell, J., "The conceptualization, measurement, and educational relevance of personal epistemology", PhD Thesis (Advisor(s): Elen, J.), (Co-Advisor(s): Clarebout, G.), 2013.
- Broeder, D., Schuurman, I., Windhouwer, M., "Experiences with the ISOcat Data Category Registry", in Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). (26-31 May 2014, Reykjavik), 2014, pp. 4565-4568.
- Broekx, L., Dreesen, K., Gemmeke, J., Van hamme, H., "Comparing and combining classifiers for self-taught vocal interfaces", in Proceedings SLPAT 2013 (21-22 August 2013, Grenoble, France), 2013, pp. 21-28.

- Broux, Y., M. Depauw, "Developing Onomastic Gazetteers and Prosopographies for the Ancient World through Named Entity Recognition and Graph Visualization: Some Examples from Trismegistos People", in: L.M. Aiello and D. McFarland (edd.), *Social Informatics. SocInfo 2014 International Workshops, GMC and Histinformatics*, Barcelona, Spain, November 10, 2014 (Lecture Notes in Computer Science 8852), Cham: Springer 2015, pp. 304-313).
- Ceci, M., Hollmen, J., Todorovski, L., Vens, C., Dzeroski, S., Eds, "Machine Learning and Knowledge Discovery in Databases. European Conference, ECML PKDD 2017 Skopje, Macedonia, September 18–22, 2017 Proceedings, Part I", "Lecture Notes in Artificial Intelligence, vol. 10534" Springer, 2017. ISBN 978-3-319-71248-2.
- Ceci, M., Hollmen, J., Todorovski, L., Vens, C., Dzeroski, S., Eds, "Machine Learning and Knowledge Discovery in Databases. European Conference, ECML PKDD 2017 Skopje, Macedonia, September 18–22, 2017 Proceedings, Part II", "Lecture Notes in Artificial Intelligence, vol. 10535" Springer, 2017. ISBN 978-3-319-71245-1.
- Cellier, P., Charnois, T., Hotho, A., Matwin, S., Moens, M., Toussaint, Y., Eds, "Interactions between data mining and natural language processing" CEUR, 2014.
- Cellier, P., Charnois, T., Hotho, A., Matwin, S., Moens, M., Toussaint, Y., Eds, "Proceedings of the workshop on interactions between data mining and natural language processing 2015 (DMNLP 2015)" CEUR, 2015.
- Cellier, P., Charnois, T., Hotho, A., Matwin, S., Moens, M., Toussaint, Y., Eds, "Proceedings of the Workshop on Interactions between Data Mining and Natural Language Processing 2016" CEUR, 2016.
- Chakravarty, J., Mirzaei, S., Tuytelaars, T., Van hamme, H., "Who's speaking? Audio-supervised classification of active speakers in video", in *Proceedings ICMI 2015* (9-13 November 2015, Seattle, Washington, USA), 2015, pp. 87-90.
- Chakravarty, J., Zegers, J., Tuytelaars, T., Van hamme, H., "Active speaker detection with audio-visual co-training", in *Proceedings ICMI 2016* (12-16 November 2016, Tokyo, Japan), 2016, pp. 312-316.
- Clarebout, G., Elen, J., Jiang, L., Lust, G., Juarez Collazo, N., "Support device usage" in *Handbook of research on educational communications and technology*, Spector, M., Merrill, D., Elen, J., Bishop, M. Eds., New York - Heidelberg - Dordrecht - London: Springer, 2014, ISBN 978-1-4614-3184-8, 978-1-4614-3185-5 (eBook), pp. 519-526.
- Clarebout, G., Elen, J., Juarez Collazo, N., Lust, G., Jiang, L., "Metacognition and the use of tools" in *International Handbook of Metacognition and Learning Technologies*, Azevedo, R., Aleven, V. Eds., New York: Springer Science+Business Media, 2013, ISBN 978-1-4419-5545-6, pp. 187-195.
- Clarebout, G., Heidig, S., "Pedagogical agents" in *Encyclopedia of the sciences of learning*, vol. 6, Seek, N. Eds., New York Dordrecht Heidelberg London: Springer, 2012, ISBN 978-1-4419-1427-9, pp. 2567-2571.
- Collell Talleda, G., Do, Q., Zhang, T., Moens, M., "Improving word representations for semantic recognition in language by adding visual contextual knowledge", (28-29 April 2016, Bern, Switzerland), 2016.
- Collell Talleda, G., Moens, M., "Is an image worth more than a thousand words? On the fine-grain semantic differences between visual and linguistic representations", in *Proceedings of the 26th International Conference on Computational Linguistics* (11-16 December 2016, Osaka, Japan), 2016, pp. 2807-2817.
- Collell Talleda, G., Moens, M., "Learning visually grounded common sense spatial knowledge for implicit spatial language", *Proceedings of the Thirteenth International Symposium on Commonsense Reasoning (Commonsense-2017)* (6-8 November 2017, University College London, UK), 2017.
- Collell Talleda, G., Zhang, T., Moens, M., "Imagined visual representations as multimodal embeddings", in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)* (4-9 February 2017, San Francisco, USA), 2017, pp. 4378-4384.
- Collell Talleda, G., Zhang, T., Moens, M., "Learning to predict: A fast re-constructive method to generate multimodal embeddings", in *Proceedings of the NIPS Workshop on Representation Learning in Artificial and Biological Neural Networks (MLINI 2016)*, 2016, art. nr. arXiv:1703.08737v1, pp. 1-9.
- Cornillie, F., "Adventures in red ink. Effectiveness of corrective feedback in digital game-based language learning", PhD Thesis (Advisor(s): Desmet, P.), (Co-Advisor(s): Clarebout, G., Van den Branden, K.), 2014.
- Cornillie, F., "Designing effective activities for focused L2 practice: insights from a game-informed experimental study", (26-30 May 2015, University of Colorado, Boulder), 2015.
- Cornillie, F., "Does instruction kill the game? Exploring learners' perceptions of corrective feedback in an immersive foreign language learning game", (14-16 June 2012, University of Notre Dame, Indiana), 2012.
- Cornillie, F., "Educationally Designed Game Environments and Feedback" in *Language, Education and Technology*, Thorne, S., May, S. Eds., Cham: Springer International Publishing, 2017, ISBN 978-3-319-02236-9, pp. 361-374.
- Cornillie, F., "Feedback in games – sleutel tot het leren van een vreemde taal", *iMinds nieuws*, 2014.
- Cornillie, F., "From drill-and-kill to the use of murder for designing effective grammar practice", (23 April 2016, 2016).

- Cornillie, F., "Great power in small games? Benefits and challenges of using educational mini-games for language practice", (13 March 2013, Kortrijk), 2013.
- Cornillie, F., "Mini-games voor taalleren: klein maar krachtig?", (11 December 2013, Gent), 2013.
- Cornillie, F., "Q&A: Frederik Cornillie on using games to teach languages", Flanders Today, 2014.
- Cornillie, F., "Worlds of wordcraft: games in het talenonderwijs", *Uit het Erasmushuis : tijdschrift van de Alumni Letteren Leuven* 5: *Ecce homo ludens*, vol. 5, 2015, pp. 137-143.
- Cornillie, F., Baten, K., De Hertog, D., "The potential of elicited imitation for oral output practice in German L2", in *CALL in a climate of change: adapting to turbulent global conditions – short papers from EUROCALL 2017* (23-26 August 2017, Southampton, UK), Borthwick, K., Bradley, L., Thouësny, S. Eds., 2017, pp. 86-91.
- Cornillie, F., Clarebout, G., Desmet, P., "Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics", *ReCALL: Journal of Eurocall*, vol. 24, no. 3, 2012, pp. 257-278.
- Cornillie, F., Clarebout, G., Desmet, P., "The role of feedback in foreign language learning through digital role playing games", in *Procedia - Social and Behavioral Sciences* (8-11 September 2010, Bordeaux), Macaire, D., Boulton, A. Eds., 2012, pp. 49-53.
- Cornillie, F., Desmet, P., "Design and empirical evaluation of controlled L2 practice through mini-games—moving beyond drill-and-kill?", (20-23 August 2014, Groningen), 2014.
- Cornillie, F., Desmet, P., "Design and empirical evaluation of controlled L2 practice through mini-games—moving beyond drill-and-kill?", in *Critical CALL – Proceedings of the 2015 EUROCALL Conference*, Padova, Italy (20-23 August 2014, Groningen), Helm, F., Bradley, L., Guarda, M., Thouësny, S. Eds., 2015, pp. 150-157.
- Cornillie, F., Desmet, P., "Mini-games for language learning" in *The Routledge Handbook of Language Learning and Technology*, Farr, F., Murray, L. Eds., Abingdon: Routledge, 2016, ISBN 9780415837873, pp. 431-445.
- Cornillie, F., Desmet, P., "Seeking out fun failure: how positive failure feedback could enhance the instructional effectiveness of CALL mini-games", (10-13 July 2013, Glasgow), 2013.
- Cornillie, F., Desmet, P., "Seeking out fun failure: how positive failure feedback could enhance the instructional effectiveness of CALL mini-games", in *Global perspectives on Computer-Assisted Language Learning. Proceedings of WorldCALL 2013* (10-13 July 2013, Glasgow), 2013, pp. 64-68.
- Cornillie, F., Fernández Castilla, B., "Using survival analysis to explain dropout in autonomous CALL practice with web-based mini-games", (16-20 May 2017, Northern Arizona University, Flagstaff, AZ), 2017.
- Cornillie, F., Grenfell, A., Windeatt, S., Desmet, P., "Challenges in specifying and evaluating a conceptual design for a task-based mini-game environment for language learning", (22-25 August 2012, Gothenburg), 2012.
- Cornillie, F., Kuhn, J., Scholz, K., "Successes, failures, and methodological issues in the design and evaluation of game-supported L2 learning environments", (6-10 May 2014, Ohio University, Athens), 2014.
- Cornillie, F., Lagatie, R., "Leveraging crowdsourced data for the automatic generation of feedback in written dialogue tasks", (12 June 2012, University of Notre Dame, Indiana), 2012.
- Cornillie, F., Lagatie, R., Desmet, P., "Language learners' use of automatically generated corrective feedback in a gamified and task-based tutorial CALL program: a pilot study", (22-25 August 2012, Gothenburg), 2012.
- Cornillie, F., Lagatie, R., Vandewaetere, M., Clarebout, G., Desmet, P., "Tools that Detectives Use: In Search of Learner-related Determinants for Usage of Optional Feedback in a Written Murder Mystery" in *Learner-Computer Interaction in Language Education*, Hubbard, P., Schulze, M., Smith, B. Eds., San Marcos, TX: CALICO, 2013, pp. 22-45.
- Cornillie, F., Thorne, S., Desmet, P., "Digital games for language learning: from hype to insight?", *ReCALL: Journal of Eurocall*, vol. 24, no. 3, 2012, pp. 243-256.
- Cornillie, F., Van den Branden, K., Desmet, P., "Effectiveness of mini-games for the development of L2 grammar knowledge: towards meaningful practice activities", (16-18 September 2015, Leuven), 2015.
- Cornillie, F., Van den Branden, K., Desmet, P., "From language play to linguistic form and back again. Lessons from an experimental study for the design of task-based language practice supported by games", in *Proceedings of XVIIth International CALL Conference. Task design and CALL*. (6-8 July 2015, Tarragona, Spain), Colpaert, J., Aerts, A., Oberhofer, M., Gutiérrez-Colón Plana, M. Eds., 2015, pp. 214-222.
- Cornillie, F., Van den Branden, K., Desmet, P., "The acquisition of English grammar through computer-assisted practice tasks and meaning-focused reading—beyond drill-and-kill?", *Eurosla 24: Book of Abstracts* (3-6 September 2014, York), 2014, pp. 63.
- Cornillie, F., Van Den Noortgate, W., Van den Branden, K., Desmet, P., "Examining focused L2 practice: from in vitro to in vivo", *Language Learning & Technology*, vol. 21, no. 1, 2017, pp. 121-145.
- Cucchiari, C., Van hamme, H., "The JASMIN speech corpus: recordings of children, non-natives and elderly people" in *Essential speech and language technology for Dutch* Springer, 2013, ISBN 978-3-642-30909-0, pp. 43-59.

- Daems J., Zenner E., Geeraerts D. (2016). Lexicale homogeniteit en lexicale voorkeur in de Nederlandse woordenschat van emoties. *Tijdschrift voor Nederlandse Taal- en Letterkunde*, 132 (4), 276-319.
- De Belder, J., "Integer Linear Programming for Natural Language Processing (Integer linear programmeren voor natuurlijke taalverwerking)", PhD Thesis (Advisor(s): Moens, M.), 2014.
- De Belder, J., de Kok, D., van Noord, G., Nauze, F., van der Beek, L., Moens, M., "Question answering of informative web pages: how summarization technology helps" in *Essential speech and language technology for Dutch: resources, tools and applications*, Spyns, P., Odijk, J. Eds., Berlin: Springer, 2012, pp. 339-357.
- De Belder, J., Moens, M., "A dataset for the evaluation of lexical simplification", in *Lecture Notes in Computer Science (11–17 March 2012, New Delhi, India)*, 2012, pp. 426-437.
- De Belder, J., Moens, M., "Coreference clustering using column generation", in *Proceedings of the 24th international conference on computational linguistics (COLING 2012) (8-15 December 2012, Mumbai, India)*, 2012, pp. 245-254.
- De Grove, F., Cornillie, F., Mechant, P., Van Looy, J., "Tapping into the field of foreign language learning games", *International Journal of Arts and Technology*, vol. 6, no. 1, 2013, pp. 22-43.
- De Hertog, D., "Distributional Models in CALL and NLP", (12-16 June 2017, Leiden), 2017.
- De Hertog, D., Cornillie, F., Desmet, P., "Context-aware Automatic Input Enhancement for Language Learners", in *CALL in context*, Colpaert, J., Aerts, A., Kern, R., Kaiser, M. Eds., 2017, pp. 209-214.
- De Hertog, D., Desmet, P., "Multiword Term Extraction through Lexical Head Selection", in *FLAIRS*, 2017, pp. 258-263.
- De Hertog, D., Desmet, P., "Structuring Usage-based Material on Syntactic and Semantic Grounds", (19-21 September 2017, Leiden), 2017.
- De Hertog, D., Heylen, K., "The prevalence of multiword term candidates in a legal corpus", in *Proceedings of the 10th Terminology and Knowledge Engineering Conference (TKE2012): New frontiers in the constructive symbiosis of terminology and knowledge engineering (19-22 June 2012, Madrid)*, Aguado de Cea, G. Eds., 2012, pp. 283-290.
- De Hertog, D., Heylen, K., Speelman, D., "Stable Lexical Marker Analysis: a corpus-based identification of lexical variation" in *Pluricentric Languages: Linguistic Variation and Sociocognitive Dimensions.*, da Silva Soares, A., Kristiansen, G. Eds. DE GRUYTER MOUTON, 2014, ISBN 978-3-11-030347-6, pp. 127-142.
- de Jong, T., Wouters, P., ter Vrugte, J., Vandercruyse, S., van Oostendorp, H., Elen, J., Verschaffel, L., Van Dooren, W., "Reflection and integration in math game-based learning for vocational students. Symposium", (27 - 31st August 2013, Munich, Germany), 2013.
- De Meester, J., Knipprath, H., Buyse, M., Ceuppens, S., De Loof, H., Goovaerts, L., Hellinckx, L., Struyf, A., Thibaut, L., Van de Velde, D., Boeve-de Pauw, J., De Cock, M., Depaepe, F., Deprez, J., Langie, G., Struyven, K., Van Petegem, P., Dehaene, W., "Leren ontwerpen van STEM-integrerend leermateriaal", *Tijdschrift voor Lerarenopleiders*, vol. 38, no. 4, 2017, art.nr. 3, pp. 27-36.
- De Mulder, W., Bethard, S., Moens, M., "A survey on the application of recurrent neural networks to statistical language modeling", *Computer Speech and Language*, vol. 30, no. 1, 2015, pp. 61-98.
- De Mulder, W., Do Thi, N., van den Broek, P., Moens, M., "Machine understanding for interactive storytelling", in *Proceedings of KICSS 2013 : 8th international conference on knowledge, information, and creativity support systems (7-9 November 2013, Kraków, Poland)*, 2013, pp. 73-80.
- De Pascale S., Marzo S., Speelman D. (2017). Evaluating regional variation in Italian: towards a change in standard language ideology?. In: Cerruti M., Crocco C., Marzo S. (Eds.), bookseries: *Language and Social Life*, vol: 6, *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*, (pp. 118-142). Berlin: De Gruyter Mouton.
- De Raedt, L., Dries, A., Thon, I., Van den Broeck, G., Verbeke, M. (2015). Inducing probabilistic relational rules from probabilistic examples. *Proceedings of 24th International Joint Conference on Artificial Intelligence (IJCAI)*. International Joint Conference on Artificial Intelligence (IJCAI). Buenos Aires, Argentina (pp. 1835-1842).
- De Smet, W., Moens, M., "Representations for multi-document event clustering", *Data Mining and Knowledge Discovery*, vol. 26, no. 3, 2013, pp. 533-558.
- Detienne, L., Raes, A., Depaepe, F. (2018). Benefits, challenges and design guidelines for synchronous hybrid learning: A systematic literature review. *EdMedia + Innovate Learning*; Amsterdam, The Netherlands.
- De Tavernier, J., Moens, M., van de Broek, P., Lavrac, N., "MUSE Ethical Report", 2012.
- De Wannemacker, S., Vandercruyse, S., Clarebout, G., Eds, "Serious Games: The Challenge", "Communications in Computer and Information Science", vol. 280", Berlin Heidelberg: Springer, 2012. ISBN 978-3-642-33814-4.
- de Wet, F., Kleynhans, N., Van Compernelle, D., Sahraeian, R., "Speech recognition for under-resourced languages: Data sharing in hidden Markov model systems", *South African Journal of Science*, vol. 113, no. 1/2, 2017, art.nr. 2016-0038.

- Degroote, H., Bischl, B., Kotthoff, L., De Causmaecker, P., "Reinforcement Learning for Automatic Online Algorithm Selection - an Empirical Study", in ITAT 2016 Proceedings (15-18 September 2016, Tatranské Matliare), 2016, pp. 93-101.
- Degroote, H., De Causmaecker, P., "Active Learning of the Best Algorithm", (28-29 January 2016, Louvain-la-Neuve), 2016.
- Degroote, H., De Causmaecker, P., "Towards a Knowledge Base for Performance Data: A Formal Model for Performance Comparison", Proceedings of the Companion Publication of the 2015 on Genetic and Evolutionary Computation Conference (11-15 July 2015, MADRID), 2015, pp. 1189-1192.
- Degroote, H., Messelis, T., De Causmaecker, P., "Improving experimental design for automatic algorithm selection techniques by using a virtual best solver", (05-06 February 2015, Antwerpen), 2015. Dehak, N., Plchot, O., Bahari, M., Burget, L., Van hamme, H., Dehak, R., "GMM weights adaptation based on subspace approaches for speaker verification", in Proceedings Odyssey 2014 (16-19 June 2014, Joensuu, Finland), 2014, pp. 48-53.
- Dekkers, G., van Waterschoot, T., Vanrumste, B., Van Den Broeck, B., Gemmeke, J., Van hamme, H., Karsmakers, P., "A multi-channel speech enhancement framework for robust NMF-based speech recognition for speech-impaired users", in INTERSPEECH 2015: proceedings (6-10 September 2015, Dresden, Germany), 2015, art. nr. 1506, pp. 746-750.
- Depauw M., B. Van Beek, "People in Greek Documentary Papyri. First Results of a Research Project", in: Journal of Juristic Papyrology 39 (2009), pp. 31-47.
- Depauw M., J. Stolk, "Linguistic Variation in Greek Papyri: Towards a New Tool for Quantitative Study", in: Greek, Roman, and Byzantine Studies 55 (2015), pp. 196-220.
- Deprez, H., Yilmaz, E., Lievens, S., Van hamme, H., "Automating speech reception threshold measurements using automatic speech recognition", in Proceedings SLPAT 2013 (August 21-22, 2013, Grenoble, France), 2013, pp. 1-6.
- Deschacht, K., De Belder, J., Moens, M., "The latent words language model", Computer Speech and Language, vol. 26, no. 5, 2012, pp. 384-409.
- Desmet, P., "L'importation du modèle historico-comparatif dans le monde francophone. Le cas d'Auguste Brachet et de Cyprien Ayer" in L'oeuvre scientifique de Cyprien Ayer (1825-1884). Grammaire, pédagogie et dialectologie, Fryba-Reber, A., Swiggers, P. Eds., Leuven - Paris - Walpole, MA: Peeters, 2013, ISBN 978-90-429-2742-1, pp. 43-76.
- Desmet, P., "La linguistique française entre 1864 et 1916: histoire de continuités et de discontinuités", HEL: Histoire, Épistémologie, Langage, vol. 34, no. 1, 2012, pp. 159-163.
- Desmet, P., "NLP-enhanced CALL: challenges and opportunities (keynote)", (11-12 Nov 2013, Stellenbosch, South-Africa), 2013.
- Desmet, P., "Seven ways to make CALL more intelligent. Towards the effective integration of NLP techniques", (6-10 May 2014, Ohio), 2014.
- Desmet, P., "The Siren CALL of Technology. Back to Pedagogical Basics (keynote)", (7-9 July 2014, Antwerp), 2014.
- Desmet, P., Cornillie, F., "Multimedia CALL" in Encyclopedia of the Sciences of Learning, Seel, N. Eds., New York: Springer, 2012, ISBN 978-1-4419-1427-9, pp. 2368-2371.
- Desmet, P., De Hertog, D., "Towards the Effective Integration of NLP in CALL (invited talk)", (16 november 2016, 2016).
- Desmet, P., Eds, "TALP et didactique", "TALP et didactique, vol. 57,3", Paris: ATALA, 2017.
- Desmet, P., Klein, J., Lamiroy, B., "Vous dites?! Répertoire d'erreurs courantes en français chez les néerlandophones", Leuven: Acco, 2012. ISBN 978-90-334-8515-2.
- Desmet, P., Montero Perez, M., "Use and effectiveness of enhanced written and audio-visual input", (10-12 May 2016, East Lansing, Michigan), 2016.
- Devriendt, J., De Causmaecker, P., Denecker, M., "Transforming constraint programs to input for local search", in The Fourteenth International Workshop on Constraint Modelling and Reformulation (31 August 2015, Cork, Ireland), 2015, pp. 1-16.
- Dirix, P., Augustinus, L., Van Eynde, F., "Een corpusonderzoek over de vervangende infinitief in het Afrikaans", (22-23 October 2015, Ghent), 2015.
- Dirix, P., Augustinus, L., Van Eynde, F., "Een toebroodje van infinitieven", (16-18 March 2017, Münster), 2017.
- Dirix, P., Augustinus, L., van Niekerk, D., Van Eynde, F., "Universal Dependencies for Afrikaans", in Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017) (22 May 2017, Gothenburg), 2017, pp. 38-47.
- Dirix, P., Van Eynde, F., "A corpus study of IPP in Afrikaans", (28-29 May 2015, Amsterdam), 2015.

- Do Thi, N., Bethard, S., Moens, M., "Adapting coreference resolution for narrative processing", in Proceedings of the 2015 conference on empirical methods in natural language processing (17-21 September 2015, Lisbon, Portugal), 2015, pp. 2262-2267.
- Do Thi, N., Bethard, S., Moens, M., "Domain adaptation in semantic role labeling using a neural language model and linguistic resources", *IEEE/ACM Transactions on Speech and Language Processing*, vol. 23, no. 11, 2015, pp. 1812-1823.
- Do Thi, N., Bethard, S., Moens, M., "Facing the most difficult case of semantic role labeling: A collaboration of word embeddings and co-training", in Proceedings of the 26th International Conference on Computational Linguistics (11-16 December 2016, Osaka, Japan), 2016, pp. 1275-1284.
- Do Thi, N., Bethard, S., Moens, M., "Improving implicit semantic role labeling by predicting semantic frame arguments", in Proceedings of the 8th International Joint Conference on Natural Language Processing (27 November - 1 December 2017, Taipei, Taiwan), 2017, pp. 90-99.
- Do Thi, N., Bethard, S., Moens, M., "Text mining for open domain semi-supervised semantic role labeling", in Proceedings of interactions between data mining and natural language processing 1st international workshop, DMNLP 2014 (15 September 2014, Nancy, France), 2014, pp. 33-48.
- Do Thi, N., Bethard, S., Moens, M., "Visualizing the content of a children's story in a virtual world: Lessons learned", in Proceedings of the EMNLP 2016 Workshop on Uphill Battles in Language Processing: Scaling Early Achievements to Robust Methods (5 November 2016, Austin, USA), 2016, pp. 39-42.
- Do Thi, N., Moens, M., "Domain adaptation in natural language understanding", Proceedings of the inter university symposium of the Belgo-Vietnamese rector's mission 2015 (19 oktober 2015, Gent, Belgium), 2015.
- Do, T., "Domain Adaptation in Natural Language Processing for Visualizing a Children's Story in a Virtual World", PhD Thesis (Advisor(s): Moens, M.), (Co-Advisor(s): Bethard, S.), 2017.
- Driesen, A., Kimmig, A., Davis, J., Belle, V., De Raedt, L., "Solving Probability Problems in Natural Language", in *Proceedings Twenty-Sixth International Joint Conference on Artificial Intelligence* (19-25 August 2017, Melbourne), 2017, pp. 3981-3987.
- Driesen, J., "Discovering Words in Speech using Matrix Factorization (Het ontdekken van woorden in spraak met behulp van matrixfactorisatie)", PhD Thesis (Advisor(s): Van hamme, H.), 2012.
- Driesen, J., Gemmeke, J., Van hamme, H., "Data-driven speech representations for NMF-based word learning", in Proceedings SAPA-SCALE (7-8 September 2012, Portland, Oregon, USA), 2012, pp. 98-103.
- Driesen, J., Gemmeke, J., Van hamme, H., "Weakly supervised keyword learning using sparse representations of speech", in Proceedings ICASSP'2012 (25-30 March 2012, Kyoto, Japan), 2012, pp. 5145-5148.
- Driesen, J., Van hamme, H., "Fast word acquisition in an NMF-based learning framework", in Proceedings ICASSP'2012 (25-30 March 2012, Kyoto, Japan), 2012, pp. 5137-5140.
- Driesen, J., Van hamme, H., "Supervised input space scaling for non-negative matrix factorization", *Signal Processing*, vol. 92, no. 8, 2012, pp. 1864-1874.
- Dusart, T., Nurani Venkitasubramanian, A., Moens, M., "Cross-modal alignment for wildlife recognition", in Proceedings of the 2nd ACM international workshop on multimedia analysis for ecological data (21 October 2013, Barcelona, Spain), 2013, pp. 9-14.
- Echeverría, V., Gomez, J., Moens, M., "Automatic labeling of forums using Bloom's taxonomy", in Proceedings of the 9th international conference on advanced data mining and applications (14-16 December 2014, Hangzhou, China), 2013, pp. 517-528.
- Eggermont, C., D'hoedt, M., "Taalportaal: een elektronisch platform voor taalvaardigheidstraining op maat", in *Van Schools tot Scriptie* (8-8 June 2012, Leiden), Smakman, D., Willemsen, L. Eds., 2012, pp. 69-79.
- Elshal, S., Mathad, M., Simm, J., Davis, J., Moreau, Y., "Topic modeling of biomedical text: from words and topics to disease and gene links", in *Proc. 2016 IEEE International Conference on Bioinformatics and Biomedicine* (Dec. 2016, Shenzhen, China), 2016, pp. 712-716.
- Elshal, S., Tranchevent, L., Sifrim, A., Ardeshirdavani, A., Davis, J., Moreau, Y., "Beegle: From literature mining to disease-gene discovery", *Nucleic Acids Research*, vol. N, no. N, 2015, pp. 1-8.
- Farnadi, G., "User Modeling in Social Media", PhD Thesis (Advisor(s): De Cock, M., Moens, M.), (Co-Advisor(s): De Cock, M.), 2017.
- Farnadi, G., Bach, S., Blondeel, M., Moens, M., De Cock, M., Getoor, L., "Statistical relational learning with soft quantifiers", in *Lecture Notes in Computer Science* (20-22 Aug 2015, Kyoto, Japan), 2016, pp. 60-75.
- Farnadi, G., Bach, S., Moens, M., Getoor, L., De Cock, M., "Extending PSL with fuzzy quantifiers", in Proceedings of StarAI 2014: Fourth International Workshop on Statistical Relational AI at AAAI 2014 (27 July, 2014, Québec, Canada), 2014, art. nr. <http://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/view/8713>.
- Farnadi, G., Bach, S., Moens, M., Getoor, L., De Cock, M., "Soft quantification in statistical relational learning", *Machine Learning*, vol. 106, no. 12, 2017, pp. 1971-1991.

- Farnadi, G., Sitaraman, G., Rohani, M., Kosinski, M., Stillwell, D., Moens, M., Davalos, S., De Cock, M., "How are you doing? Emotions and personality in Facebook", in Proceedings of the EMPIRE Workshop of the 22nd International Conference on User Modeling, Adaptation and Personalization (UMAP 2014) (11 July 2014, Aalborg, Denmark), 2014, art. nr. http://ceur-ws.org/Vol-1181/empire2014_paper_05.pdf.
- Farnadi, G., Sitaraman, G., Sushmita, S., Celli, F., Kosinski, M., Stillwell, D., Davalos, S., Moens, M., De Cock, M., "Computational personality recognition in social media", *User Modeling and User-adapted Interaction*, vol. 26, no. 2-3, 2016, art.nr. DOI 10.1007/s11257-016-9171-0, pp. 109-142.
- Farnadi, G., Zeinab, M., Ivan, K., Jacob, N., Ankur, T., Moens, M., De Cock, M., "Scalable adaptive label propagation in Grappa", in Proceedings of the 2015 IEEE International Conference on Big Data (Big Data), 2015, pp. 1485-1491.
- Farnadi, G., Zoghbi, S., Moens, M., De Cock, M., "Crawling a ground truth dataset for ad personalization", Proceedings of the 5th AUGent workshop on language technology and computational intelligence, (Gent), 2012, pp. 7.
- Farnadi, G., Zoghbi, S., Moens, M., De Cock, M., "How well do your Facebook status updates express your personality?", Proceedings of the 22nd edition of the annual Belgian-Dutch conference on machine learning (BENELEARN), 2013, pp. 88.
- Farnadi, G., Zoghbi, S., Moens, M., De Cock, M., "Recognising personality traits using facebook status updates", in Proceedings of the workshop on computational personality recognition (WCPR13) at the 7th international AAI conference on weblogs and social media (ICWSM13) (11 July 2013, Boston, USA), 2013, art. nr. <http://mypersonality.org/wiki/doku.php?id=wcpr13>.
- Feremans, L., Cule, B., Vens, C., Goethals, B., "Combining Instance and Feature neighbors for Efficient Multi-label Classification", in Proceedings of IEEE International Conference on Data Science and Advanced Analytics (19-21 October 2017, Tokyo, Japan), 2017, pp. 109-118.
- Ferro, N., Crestani, F., Moens, M., Mothe, J., Silvestri, F., Di Nunzio, G., Hauff, C., Silvello, G., Eds, "Advances in Information Retrieval - Proceedings of the 38th European Conference on IR Research, ECIR 2016", "Lecture Notes in Computer Science, vol. 9626" Springer, 2016. ISBN 978-3-319-30670-4.
- Fievez, I., Montero Perez, M., Cornillie, F., Desmet, P., "Vocabulary acquisition from audio-visual input: from narrow to extensive viewing", (3-4 July 2017, University of Reading), 2017.
- François, T., Volodina, E., Pilán, I., Tack, A., "SVALex: a CEFR-graded Lexical Resource for Swedish Foreign and Second Language Learners", in Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016) (23-28 May 2016, Portorož, Slovenia), Calzolari, N., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S. Eds., 2016, pp. 213-219.
- Frasconi, P., Costa, F., De Raedt, L., De Grave, K. (2014). kLog: A language for logical and relational learning with kernels. *Artificial Intelligence*, 217, 117-143.
- Geeraerts D. (2016). Các lí thuyết ngữ nghĩa học từ vựng [Vietnamese translation of 2010 Theories of Lexical Semantics]. Hanoi: Nhà Xuất Bản Đại Học Quốc Gia Hà Nội.
- Geeraerts D. (2016). Entrenchment as onomasiological salience. In: Schmid H. (Eds.), bookseries: Language and the Human Lifespan (LHLS), Entrenchment and the psychology of language learning: how we reorganize and adapt linguistic knowledge, (pp. 153-174). Berlin, Boston: De Gruyter Mouton.
- Geeraerts D. (2016). The sociosemiotic commitment. *Cognitive Linguistics*, 27 (4), 527-542.
- Geeraerts D. (2017). Conceptual Structure and Conceptual Variation. Shanghai: Shanghai Foreign Language Education Press.
- Geeraerts D. (2017). Distributionalism, old and new. In: Makarova A., Dickey S. M., Divjak, D. (Eds.), Each Venture a New Beginning. Studies in Honor of Laura A. Janda, (pp. 29-38). Bloomington: Indiana: Slavica.
- Geiger, J., Weninger, F., Hurmalainen, A., Gemmeke, J., Wöllmer, M., Schuller, B., Rigoll, G., Virtanen, T., "The TUM+TUT+KUL approach to the CHiME challenge 2013: Multi-stream ASR exploiting BLSTM networks and sparse NMF", in Proceedings CHiME 2013 (1 June 2013, Vancouver, Canada), 2013, pp. 25-30.
- Gemmeke, J., "Resolving sparse representation model mismatches through linear transforms", (8-11 July 2013, Lausanne, Switzerland), 2013.
- Gemmeke, J., Hurmalainen, A., Virtanen, T., "HMM-regularization for NMF-based noise robust ASR", in Proceedings CHiME 2013 (1 June 2013, Vancouver, Canada), 2013, pp. 47-52.
- Gemmeke, J., Ons, B., Tessema, N., Van hamme, H., Van de Loo, J., De Pauw, G., Daelemans, W., Huyghe, J., Derboven, J., Vuegen, L., Van Den Broeck, B., Karsmakers, P., Vanrumste, B., "Self-taught assistive vocal interfaces: An overview of the ALADIN project", in Proceedings Interspeech 2013 (25-29 August 2013, Lyon, France), Bimbot, F., Fougeron, C., Pellegrino, F. Eds., 2013, art. nr. 760, pp. 2038-2043.

- Gemmeke, J., Remes, U., "Missing-data techniques: Feature reconstruction" in Techniques for noise robustness in automatic speech recognition, Virtanen, T., Singh, R., Raj, B. Eds. Wiley, 2012, ISBN 978-1-1199-7088-0, pp. 399-432.
- Gemmeke, J., Sehgal, S., Cunningham, S., Van hamme, H., "Dysarthric vocal interfaces with minimal training data", in Proceedings SLT 2014 (7-10 December 2014, South Lake Tahoe, Nevada, USA), 2014, pp. 1-6.
- Gemmeke, J., van de Loo, J., De Pauw, G., Driesen, J., Van hamme, H., Daelemans, W., "A self-learning assistive vocal interface based on vocabulary learning and grammar induction", in Proceedings Interspeech 2012 (9-13 September 2012, Portland, Oregon, USA), 2012, pp. 1-4.
- Gemmeke, J., Van hamme, H., "Advances in noise robust digit recognition using hybrid exemplar-based techniques", in Proceedings Interspeech 2012 (9-13 September 2012, Portland, Oregon, USA), 2012, pp. 1-4.
- Gemmeke, J., Virtanen, T., Demuynck, K., "Exemplar-based joint channel and noise compensation", in Proceedings ICASSP'2013 (26-31 May 2013, Vancouver, Canada), 2013, pp. 868-872.
- Gemmeke, J., Vuegen, L., Karsmakers, P., Vanrumste, B., Van hamme, H., "An exemplar-based NMF approach to audio event detection", in Proceedings WASPAA 2013 (20-23 October 2013, New Paltz, New York, USA), Gannot, S., Habets, E. Eds., 2013, pp. 1-4.
- Gizem Cinar, Y., Zoghbi, S., Moens, M., "Inferring user interests on social media from text and images", in Proceedings of SoMeRA 2015: 2nd international workshop on social media retrieval and analysis at ICDM 2015 (14 November 2015, Atlantic City, NJ, USA), 2015.
- Glavaš, G., Šnajder, J., Kordjamshidi, P., Moens, M., "HiEve: A corpus for extracting event hierarchies from news stories", in Proceedings of 9th language resources and evaluation conference (26-31 May 2014, Reykjavik, Iceland), 2014, pp. 3678-3683.
- Gomez, J., Boiy, E., Moens, M., "Highly discriminative statistical features for email classification", Knowledge and Information Systems, vol. 31, no. 1, 2012, pp. 23-53.
- Gomez, J., Hoskens, S., Moens, M., "Evolutionary learning of meta-rules for text classification", Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '17) (15-19 July 2017, Berlin, Germany), 2017, pp. 131-132.
- Gomez, J., Moens, M., "A survey of automated hierarchical classification of patents" in Professional search in the modern world, Paltoglou, G., Loizides, F., Preben, H. Eds., Berlin, Germany: Springer, 2014, ISBN 978-3-319-12511-4, pp. 215-249.
- Gomez, J., Moens, M., "Document categorization based on minimum loss of reconstruction information", in Lecture Notes in Computer Science (27 October - 4 November 2012, San Luis Potosí, Mexico), 2012, pp. 91-103.
- Gomez, J., Moens, M., "Hierarchical classification of web documents by stratified discriminant analysis", in Lecture Notes in Computer Science (2-3 July 2012, Vienna, Austria), 2012, pp. 94-108.
- Gomez, J., Moens, M., "Minimizer of the reconstruction error for multi-class document categorization", Expert Systems with Applications, vol. 41, no. 3, 2014, pp. 779-930.
- Gomez, J., Moens, M., "Multilayered class discrimination in large-scale taxonomies", in Frontiers in Artificial Intelligence (10-12 September 2012, San Sebastian, Spain), 2012, pp. 615-625.
- Gomez, J., Moens, M., "PCA document reconstruction for email classification", Computational Statistics & Data Analysis, vol. 56, no. 3, 2012, pp. 741-751.
- Gomez, J., Tommassi, T., Zoghbi, S., Moens, M., "What would they say? Predicting user's comments in Pinterest", IEEE Latin America Transactions, vol. 14, no. 4, 2016, pp. 2013-2019.
- Gosh, S., Ghosh, K., Chakraborty, T., Ganguly, D., Jones, G., Moens, M., Eds, "Proceedings of the First International Workshop on Exploitation of Social Media for Emergency Relief and Preparedness" CEUR, 2017.
- Gyllstrom, K., Eickhoff, C., de Vries, A., Moens, M., "The downside of markup: examining the harmful effects of CSS and javascript on indexing today's web", in Proceedings of the 21st ACM international conference on information and knowledge management (CIKM 2012) (29 October - 2 November 2012, Maui, Hawaii), 2012, pp. 1990-1994.
- Gyllstrom, K., Moens, M., "Surfin' Wikipedia: an analysis of the Wikipedia (non-random) surfer's behavior from aggregate access data", in Proceedings of the fourth information interaction in context symposium (IIX 2012) (21-24 August 2012, Nijmegen), 2012, pp. 155-163.
- Haesen, M., Meskens, J., Luyten, K., Coninx, K., Becker, J., Tuytelaars, T., Poulisse, G., Pham The, P., Moens, M., "Finding a needle in a haystack: an interactive video archive explorer for professional video searchers", Multimedia Tools and Applications, vol. 63, no. 2, 2013, pp. 331-356.
- Hens, K., Do Thi, N., Moens, M., "Automatic conversion of an episodes script to a 3D movie", Proceedings of the 25th meeting of computational linguistics in the Netherlands (5-6 February 2015, Antwerp), 2015, pp. 80.
- Heylen K., Bertels A. (2016). Sémantique distributionnelle en linguistique de corpus. Langages, 2016 (201), 51-64.

- Heylen, K., Bond, S., De Hertog, D., Vulic, I. and Kockaert, H. "TermWise: A CAT-tool with Context-Sensitive Terminological Support". In Proceedings of the 9th Language Resources and Evaluation Conference. 2014
- Heylen, K.; Bond, S.; De Hertog, D.; Kockaert, H. ; Steurs, F.; Vulić, I (2014) "TermWise: Leveraging Big Data for Terminological Support" in Legal Translation Proceedings of the 11th Terminology and Knowledge Engineering Conference: Ontology, Terminology & Text Mining (TKE), p. 167-176.
- Heylen, K.; De Hertog, D. (2015) "Automatic Term Extraction" In: Kockaert, Hendrik J.; Steurs, Frieda (Ed.), Handbook of Terminology: Volume 1. Amsterdam: John Benjamins Publishing Company, p. 199-219.
- Heylen, K.; Steurs, F. (2014), Translating legal and administrative language: How to deal with legal terms and their flexible meaning potential, *Turjuman. Journal of Translation Studies* 23 (2), p. 96-146.
- Heylen, K.; Wielfaert, T.; Speelman, D.; Geeraerts, D. (2015), "Monitoring Polysemy. Word Space Models as a Tool for Large-Scale Lexical Semantic Analysis", *Lingua* 157 , p. 153-172.
- Heyman, G., Vulić, I., Moens, M., "Bilingual Lexicon Induction by Learning to Combine Word-Level and Character-Level Representations", Book of abstracts CLIN27, 2017, pp. 66-67.
- Heyman, G., Vulić, I., Moens, M., "Bilingual Lexicon Induction by Learning to Combine Word-Level and Character-Level Representations", in Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (3-7 April 2017, Valencia, Spain), 2017, pp. 1085-1095.
- Heyman, G., Vulic, I., Moens, M., "C-BiLDA extracting cross-lingual topics from non-parallel texts by distinguishing shared from unshared content", *Data Mining and Knowledge Discovery*, vol. 30, no. 5, 2016, art.nr. 10.1007/s10618-015-0442-x, pp. 1299-1323.
- Hilson, J., Besnehard, A., Desmet, P., "The technological concept behind the Franel language learning environment: Recent developments and outlook", in *Procedia - Social and Behavioral Sciences* (8-11 September 2010, Bordeaux), Macaire, D., Boulton, A. Eds., 2012, pp. 100-105.
- Hurmalaianen, A., Gemmeke, J., Virtanen, T., "Compact long context spectral factorisation models for noise robust recognition of medium vocabulary speech", in Proceedings CHI ME 2013 (1 June 2013, Vancouver, Canada), 2013, pp. 13-18.
- Hurmalaianen, A., Gemmeke, J., Virtanen, T., "Detection, separation and recognition of speech from continuous signals using spectral factorisation", in Proceedings EUSIPCO 2012 (27-31 August 2012, Bucharest, Romania), 2012, pp. 2649-2653.
- Hurmalaianen, A., Gemmeke, J., Virtanen, T., "Modelling non-stationary noise with spectral factorisation in automatic speech recognition", *Computer Speech and Language*, vol. 27, no. 3, 2013, pp. 763-779.
- Jacobs, G., Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Word Sense Disambiguation in Text-to-Pictograph Translation", Book of abstracts CLIN26 (18 December 2015, Amsterdam), 2015.
- Jacques, I., De Causmaecker, P., "Automated learning topic refinement of exercises based on logging data", in INTED2012 Proceedings (March 2012, Spain), 2012.
- Jans, B., Bethard, S., Vulic, I., Moens, M., "Skip n-grams and ranking functions for predicting script events", in Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012) (23-27 April 2012, Avignon, France), 2012, pp. 336-344.
- Jia, X., Gavves, S., Fernando, B., Tuytelaars, T., "Guiding long-short term memory for image caption generation", in Proceedings of the International Conference on Computer Vision (ICCV), 2015.
- Juarez Collazo, N., "Unravelling tool usage. Analyzing the functionality of tools and the variables influencing the use of tools in computer-based learning environments", PhD Thesis (Advisor(s): Elen, J.), (Co-Advisor(s): Clarebout, G.), 2014.
- Juarez Collazo, N., Corradi, D., Elen, J., Clarebout, G., "Tool use of experienced learners in computer-based learning environments: Can tools be beneficial?", *Higher Education Studies*, vol. 4, no. 1, 2014, pp. 26-42.
- Juarez Collazo, N., Elen, J., Clarebout, G., "Effects of prior knowledge, perceptions and explained tool functionality on tool use and performance", Book of Abstracts 15th Biennial Conference EARLI 2013 (27-31 August 2013, Munich, Germany), 2013, pp. 315.
- Juarez Collazo, N., Elen, J., Clarebout, G., "Learners, tools, and interventions: The complexity of tool use in computer-based learning environments", (27 April - 1 May 2013, San Francisco, CA, USA), 2013.
- Juarez Collazo, N., Elen, J., Clarebout, G., "Perceptions for tool use: In search of a tool use model", in Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2012 (26-29 June 2012, Denver, CO, USA), Amiel, T., Wilson, B. Eds., 2012, art. nr. 35878, pp. 2905-2912.
- Juarez Collazo, N., Elen, J., Clarebout, G., "The multiple effects of combined tools in computer-based learning environments", *Computers in Human Behavior*, vol. 51, no. Part A, 2015, pp. 82-95.
- Juarez Collazo, N., Elen, J., Clarebout, G., "The toll of (not) using tools in computer learning environments and the role of self-efficacy", in Canada International Conference on Education (CICE-2012) (18-21 June 2012, Ontario, Canada), Shoniregun, C., Akmayeva, G. Eds., 2012, pp. 721-725.

- Juarez Collazo, N., Elen, J., Clarebout, G., "The unraveling of tool use: Answering the unanswered?", *Journal of Global Research in Computer Science*, vol. 6, no. 7, 2015, pp. 1-10.
- Juarez Collazo, N., Elen, J., Clarebout, G., "Tool use in computer learning environments: Enhancing learning?", (11-13 September 2012, Bari, Italy), 2012.
- Juarez Collazo, N., Jiang, L., Elen, J., Clarebout, G., "Tool use and performance: Relationships between tool- and learner-related characteristics in a computer-based learning environment", *TOJET*, vol. 2, no. 2, 2013, pp. 330-345.
- Juarez Collazo, N., Wu, X., Elen, J., Clarebout, G., "Tool use in computer-based learning environments: Adopting and extending the technology acceptance model", *ISRN Education*, vol. 2014, 2014, art.nr. ID 736931.
- Juarez Collazo, S., Elen, J., Clarebout, G., "The role of learner variables, tool variables and interventions on tool use and performance in computer learning environments", in *ICERI 2012 Proceedings (19-21 November 2012, Madrid, Spain)*, Gómez Chova, L., López Martínez, A., Candel Torres, I. Eds., 2012, art. nr. 1013, pp. 48-57.
- Kadengye, D., "Explanatory Item Response Modeling of Structurally Involved Educational Data. With Applications to Electronic Learning Environments", PhD Thesis (Advisor(s): Van Den Noortgate, W.), (Co-Advisor(s): Ceulemans, E.), 2013.
- Kadengye, D., Ceulemans, E., Van Den Noortgate, W., "A generalized longitudinal mixture IRT model for measuring differential growth in learning environments", *Behavior Research Methods*, vol. 46, 2014, pp. 823-840.
- Kadengye, D., Ceulemans, E., Van Den Noortgate, W., "A mixture IRT growth model for longitudinal data from e-learning environments", (15 - 17 May 2013, Clane, Kildare, Ireland), 2013.
- Kadengye, D., Ceulemans, E., Van Den Noortgate, W., "Direct likelihood analysis and multiple imputation for missing item scores in multilevel cross-classification educational data", *Applied Psychological Measurement*, vol. 38, no. 1, 2014, pp. 61-80.
- Kadengye, D., Ceulemans, E., Van Den Noortgate, W., "Missing item scores in multilevel cross-classified educational data: An empirical Bayes Multiple Imputation approach", (22-23 May 2012, Connecticut, USA), 2012.
- Kadengye, D., Ceulemans, E., Van Den Noortgate, W., "Modeling growth in electronic learning environments using a longitudinal random item response model", *Journal of Experimental Education*, vol. 83, no. 2, 2015, pp. 175-202.
- Kadengye, D., Ceulemans, E., Van Den Noortgate, W., "Multiple imputation for missing binary item scores in multilevel cross-classified educational data when the Analysis and Imputation models differ", (7-8 June 2012, 2012).
- Kadengye, D., Cools, W., Ceulemans, E., Van Den Noortgate, W., "Simple Imputation Methods versus Direct Likelihood Analysis for Missing Item Scores in Multilevel Educational Data", *Behavior Research Methods*, vol. 44, no. 2, 2012, pp. 516-531.
- Kadengye, D., Van Den Noortgate, W., Ceulemans, E., "Application of Random Item IRT Models to Longitudinal Data from Electronic Learning Environments", (15 - 17 May 2013, Clane, Kildare, Ireland), 2013.
- Kakouros, S., Pelemans, J., Verwimp, L., Wambacq, P., Räsänen, O., "Analyzing the contribution of top-down lexical and bottom-up acoustic cues in the detection of sentence prominence", in *Proceedings Interspeech 2016 (8-12 September 2016, San Francisco, California, USA)*, 2016, art. nr. 926, pp. 1074-1078.
- Karan, M., Glavaš, G., Šnajder, J., Dalbelo Bašić, B., Vulic, I., Moens, M., "TKLBIIR: detecting twitter paraphrases with TweetingJay", in *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015) (4-5 June 2015, Denver, CO, USA)*, 2015, pp. 70-74.
- Kemps-Snijders, M., Schuurman, I., Daelemans, W., Demuyne, K., Desplanques, B., Hoste, V., Huijbrechts, M., Martens, J., Paulussen, H., Pelemans, J., Reynaert, M., Vandeghinste, V., van den Bosch, A., van den Heuvel, H., van Gompel, M., van Noord, G., Wambacq, P., "TTNWW to the Rescue: No Need to Know How to Handle Tools and Resources" in *CLARIN in the Low Countries*, Odijk, J., van Hessen, A. Eds., London: Ubiquity Press, 2017, ISBN 978-1-911529-24-8, 978-1-911529-25-5, 978-1-911529-26-2, 978-1-911529-27-9, pp. 83-93.
- Kerckhof, L., Seurnyck, Y., Haspeslagh, S., Vanderplancke, T., "A Coherent LMS for E-Learning within the Care Sector", in *Proceedings of IAC 2017 in Bratislava (7-8 July 2017, Bratislava)*, 2017, pp. 17-23.
- Keronen, S., Kallasjoki, H., Remes, U., Brown, G., Gemmeke, J., Palomäki, K., "Mask estimation and imputation methods for missing data speech recognition in a multisource reverberant environment", *Computer Speech and Language*, vol. 27, no. 3, 2013, pp. 798-819.
- Kocev, D., Vens, C., Struyf, J., Dzeroski, S., "Tree ensembles for predicting structured output", *Pattern Recognition*, vol. 46, no. 3, 2013, pp. 817-833.
- Kolomiyets, O., "Algorithms for Temporal Information Processing of Text and their Applications (Algoritmen voor het verwerken van temporele informatie in tekst en hun toepassingen)", PhD Thesis (Advisor(s): Moens, M., De Schreye, D.), 2012.

- Kolomiyets, O., Bethard, S., Moens, M., "Extracting narrative timelines as temporal dependency structures", in Proceedings of the 50th annual meeting of the association for computational linguistics (8-14 July 2012, Jeju, Republic of Korea), 2012, pp. 88-97.
- Kolomiyets, O., Kordjamshidi, P., Bethard, S., Moens, M., "SemEval-2013 task 3: Spatial role labeling", in Second joint conference on lexical and computational semantics (*SEM), Volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013) (14-15 June 2013, Atlanta, USA), 2013, pp. 255-266.
- Kolomiyets, O., Moens, M., "KUL: A data-driven approach to temporal parsing of documents", in Proceedings of the second joint conference on lexical and computational semantics (*SEM), Volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013) (14-15 June 2013, Atlanta, USA), 2013, pp. 83-87.
- Kolomiyets, O., Moens, M., "Towards animated visualization of actors and actions in a learning environment", in Proceedings of the 3rd international workshop on evidence based and user centred technology enhanced learning (EbuTEL 2013) (16 September 2013, Trento, Italy), 2014, pp. 45-53.
- Kordjamshidi, P., "Structured Machine Learning for Mapping Natural Language to Spatial Ontologies (Gestructureerd machinaal leren voor het omzetten van natuurlijke taal naar ruimtelijke ontologieën)", PhD Thesis (Advisor(s): Moens, M.), 2013.
- Kordjamshidi, P., Bethard, S., Moens, M., "SemEval-2012 task 3: Spatial role labeling", in { *SEM 2012 } : The First Joint Conference on Lexical and Computational Semantics -- Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation { (SemEval 2012) } (7-8 June, Montreal- Canada), 2012, pp. 365-373.
- Kordjamshidi, P., Frasconi, P., van Otterlo, M., Moens, M., De Raedt, L., "Relational learning for spatial relation extraction from natural language", in Lecture notes in artificial intelligence (31st July-3rd August, Windsor), Muggleton, S., Tamaddoni-Nezhad, A., Lisi, F. Eds., 2012, pp. 204-220.
- Kordjamshidi, P., Hois, J., van Otterlo, M., Moens, M., "Learning to interpret spatial natural language in terms of qualitative spatial relations" in Representing space in cognition: interrelations of behavior, language, and formal models, Tenbrink, T., Wiener, J., Claramunt, C. Eds. Oxford University Press, 2013, ISBN 9780199679911, pp. 115-146.
- Kordjamshidi, P., Massa, W., Provoost, T., Moens, M., "Machine reading for extraction of bacteria and habitat taxonomies", in Lecture Notes in Computer Science, 2015, pp. 239-255.
- Kordjamshidi, P., Moens, M., "Designing constructive machine learning models based on generalized linear learning techniques", in Proceedings of the NIPS workshop on constructive machine learning (10 December 2013, Lake Tahoe, Nevada, USA), 2013, pp. 1-5.
- Kordjamshidi, P., Moens, M., "Global machine learning for spatial ontology population", Journal of Web Semantics, vol. 30, 2015, pp. 3-21.
- Kordjamshidi, P., Moens, M., "Spatial role labeling using structured support vector machines", Proceedings of 21st Belgian-Dutch conference on machine learning (BeneLearn) (24-25 May, Ghent), 2012, pp. 71-71.
- Kordjamshidi, P., Moens, M., "Structured machine learning for mapping natural language to spatial ontologies", (8-19 July 2013, Leuven), 2013.
- Kordjamshidi, P., Rahgooy, T., Moens, M., Pustejovsky, J., Manzoor, U., Roberts, K., "CLEF 2017: Multimodal spatial role labeling (mSpRL) task overview", Lecture Notes in Computer Science, vol. 10456, 2017, pp. 367-376.
- Kordjamshidi, P., Roth, D., Moens, M., "Structured learning for spatial information extraction from biomedical text: bacteria biotopes", BMC Bioinformatics, vol. 16, 2015, art.nr. 129, pp. 129.
- Kordjamshidi, P., van Otterlo, M., Moens, M., "Spatial role labeling annotation scheme" in Handbook of Linguistic Annotation, Pustejovsky, J., Ide, N. Eds. Springer Verlag, 2017, ISBN 978-94-024-0879-9, pp. 1025-1052.
- Kotzé, G., Vandeghinste, V., Martens, S., Tiedemann, J., "Large Aligned Treebanks for Syntax-based Machine Translation", Language Resources and Evaluation, vol. 51, no. 2, 2016, art.nr. DOI 10.1007/s10579-016-9369-0, pp. 249-279.
- Kovács, G., Tóth, L., Van Compernelle, D., "Selection and enhancement of Gabor filters for automatic speech recognition", International Journal of Speech Technology, vol. 18, no. 1, 2015, pp. 1-16.
- Laenen, K., Zoghbi, S., Moens, M., "Cross-modal search for fashion attributes", in Proceedings of the KDD 2017 Workshop on Machine Learning Meets Fashion (14 August 2017, Halifax, Canada), 2017, pp. 1-10.
- Leeuwenberg, T., Moens, M., "KU Leuven-LIIR at SemEval 2016 Task 12: Detecting Narrative Containment in Clinical Records", in Proceedings of SemEval 2016, San Diego, California), 2016, pp. 1280-1285.
- Leeuwenberg, T., Moens, M., "KU Leuven-LIIR at SemEval-2017 task 12: Cross-domain temporal information extraction from clinical records", in Proceedings of SemEval-2017 International Workshop on Semantic Evaluation (3-4 August 2017, Vancouver, Canada), 2017, pp. 1030-1034.

- Leeuwenberg, T., Moens, M., "Structured learning for temporal relation extraction from clinical records", in Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (3-7 April 2017, Valencia, Spain), 2017, pp. 1150-1158.
- Levshina, N.; Heylen, K. (2014) "A radically data-driven Construction Grammar: Experiments with Dutch causative constructions", In: Boogaart, R.; Colleman, T.; Rutten, G. (Ed.), *Extending the Scope of Construction Grammar*. Berlin: Mouton de Gruyter, p. 17-46.
- Luca, S., Vuegen, L., Van hamme, H., Karsmakers, P., Vanrumste, B., "Decision support systems for home monitoring applications: Classification of activities of daily living and epileptic seizures" in *Machine Learning for Healthcare Technologies*, Clifton, D. Eds., London: IET, Institution of Engineering and Technology, 2016, ISBN 9781849199780, 9781849199797, pp. 271-291.
- Ludwig, O., Do Thi, N., Moens, M., "Mapping from written stories to virtual reality", Proceedings of the 25th meeting of computational linguistics in the Netherlands (5-6 February 2015, Antwerp), 2015.
- Ludwig, O., Do Thi, N., Smith, C., Cavazza, M., Moens, M., "Learning to extract action descriptions from narrative text", *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 2017, 2017, pp. 1-15.
- Lust, G., "Opening the Black Box. Students' Tool-use within a Technology-Enhanced learning environment: An Ecological-Valid Approach", PhD Thesis (Advisor(s): Elen, J.), (Co-Advisor(s): Clarebout, G.), 2012.
- Lust, G., Elen, J., Clarebout, G., "Adopting web-casts over time: The influence of perceptions and attitudes", *Journal of Computing in Higher Education*, vol. 24, no. 1, 2012, pp. 40-57.
- Lust, G., Elen, J., Clarebout, G., "Het gebruik van webcolleges: wetenschappelijk perspectief", (23 oktober 2012, Katho, Kortrijk), 2012.
- Lust, G., Elen, J., Clarebout, G., "Measuring students' strategy-use within a CMS supported course through students' tool-use patterns", *Book of Abstracts 15th Biennial Conference EARLI 2013 (27-31 August 2013, Munich, Germany)*, 2013, pp. 571-572.
- Lust, G., Elen, J., Clarebout, G., "Online measurement perspectives for students' strategy-use: Tool-use within a content management system", *Education Research International*, vol. 2012, no. 2012, 2012, art.nr. 702315, pp. 1-11.
- Lust, G., Elen, J., Clarebout, G., "Regulation of tool-use within a blended course: Student differences and performance effects", *Computers and Education*, vol. 60, no. 1, 2013, pp. 385-395.
- Lust, G., Elen, J., Clarebout, G., "Regulation of tool-use within a blended learning environment", (11-13 September 2012, Bari, Italy), 2012.
- Lust, G., Elen, J., Clarebout, G., "Students' learning within a content management-supported learning environment: Is timing everything?", *Book of Abstracts 15th Biennial Conference EARLI 2013 (27-31 August 2013, Munich, Germany)*, 2013, pp. 293.
- Lust, G., Elen, J., Clarebout, G., "Students' study strategies within a Content Management System: Behavioral indicators", (18-21 September, 2012, Cadiz, Spain), 2012.
- Lust, G., Elen, J., Clarebout, G., "Students' tool-use within a web enhanced course: Explanatory mechanisms of students' tool-use pattern", *Computers in Human Behavior*, vol. 29, no. 5, 2013, pp. 2013-2021.
- Lust, G., Juarez Collazo, N., Elen, J., Clarebout, G., "Content Management Systems: Enriched learning opportunities for all?", *Computers in Human Behavior*, vol. 28, no. 3, 2012, pp. 795-808.
- Lust, G., Vandewaetere, M., Elen, J., Clarebout, G., "Students' compliance to technology-supported learning environments: Tool-use perspective" in *Learning environments: Technologies, challenges and impact assessment*, Atkinson, R. Eds. Nova Science Publishers, Inc., 2013, ISBN 1628085940.
- Lust, G., Vandewaetere, M., Elen, J., Clarebout, G., "Tool-use in a content management system: A matter of timing?", *Learning Environments Research*, vol. 17, no. 3, 2012, pp. 319-337. Maertens, M., Castelein, N., Vincke, L., "Assistive technology & communication problems", (4-7 September 2012, Saint-Andrews, Schotland, UK), 2012.
- Maertens, M., "Importing pen-and-paper content into an educational math game: a transfer with added difficulty", (27-29 August 2014, Rotterdam, the Netherlands), 2014.
- Maertens, M., Castelein, N., Vincke, L., "Assistive technology & communication problems", (4-7 September 2012, Saint-Andrews, Schotland, UK), 2012.
- Maertens, M., Cornillie, F., Desmet, P., Vandewaetere, M., "Adaptive and adaptable gameplay for children: challenges and pitfalls", (12-15 June 2013, Oslo (Sweden)), 2013.
- Maertens, M., Cornillie, F., Desmet, P., Vandewaetere, M., "Adaptive and adaptable gameplay for children: challenges and pitfalls", (27 April - 2 May 2013, Paris, France), 2013.
- Maertens, M., Cornillie, F., Desmet, P., Vandewaetere, M., "Adaptive and Adaptable Gameplay: an Instructional Design Approach for Individual Differences", (12-15 June 2013, Oslo, Sweden), 2013.
- Maertens, M., Desmet, B., Defrenne, C., "Report: empowerment vision and theoretical foundation", 2015.

- Maertens, M., Desmet, P., Cornillie, F., Vandewaetere, M., "It's all in the game: Assessing the difficulty of math exercises in an educational game", (27 - 31st August 2013, Munich, Germany), 2013.
- Maertens, M., Desmet, P., Cornillie, F., Vandewaetere, M., "The added value of adaptivity for providing challenge in educational math games", (27 - 31st August 2013, Munich, Germany), 2013.
- Maertens, M., Faifer, D., Clarebout, G., "Bachelor Degree in Law at KU Leuven-Kulak. A Blended Learning Program", (3-6 July 2012, Leuven, Belgium), 2012.
- Maertens, M., Van Den Noortgate, W., Ceulemans, E., Elen, J., Thys, A., Vandewaetere, M., Van Dooren, W., "Thematic poster: Stat+: An enrichment of a learning environment for statistics", EARLI 2015, Book of Abstracts (25-29 August, 2015, Limassol, Cyprus), 2015, pp. 666-667.
- Maertens, M., Vandewaetere, M., Cornillie, F., Desmet, P., "From pen-and-paper content to educational math game content for children: a transfer with added difficulty", *International Journal of Child - Computer Interaction*, vol. 2, no. 2, 2014, pp. 85-92.
- Maertens, M., Vandewaetere, M., Cornillie, F., Desmet, P., "Performance in educational math games: Is it a question of math knowledge?" in *Describing and studying domain-specific serious games*, Torbeyns, J., Lehtinen, E., Elen, J. Eds., Cham - Heidelberg - New York - Dordrecht - London: Springer, 2015, ISBN 978-3-319-20275-4, 978-3-319-20276-1, pp. 117-131.
- Maertens, M., Vandewaetere, M., Cornillie, F., Desmet, P., "Performance in educational math games: Is it a question of math skills?", (15-17 October 2014, Belgium: Leuven), 2014. Marquardt, J., Farnadi, G., Vasudevan, G., Moens, M., Davalos, S., Teredesai, A., De Cock, M., "Age and gender identification in social media", in *Proceedings of CLEF 2014 Evaluation Labs (15-18 September, 2014, Sheffield, UK)*, 2014, pp. 1129-1136.
- Massa, W., Kordjamshidi, P., Provoost, T., Moens, M., "Machine reading of biological texts: bacteria-biotope extraction", in *Proceedings of the 6th international conference on bioinformatics models, methods and algorithms (12-15 January 2015, Lisbon, Portugal)*, 2015, pp. 55-64.
- Medina Maestro, J., Saggion, H., Schuurman, I., Sevens, L., O'Flaherty, J., De Vlieghe, A., Daems, J., "Towards Integrating People with Intellectual Disabilities in the Digital World", in *Proceedings of the 7th International Workshop on Intelligent Environments Supporting Healthcare and Well-being (12-13 September 2016, Middlesex (London, UK))*, 2016, pp. 348-357.
- Mirzaei, S., "Application of Bayesian Matrix Factorization Methods in Audio Signal Processing", PhD Thesis (Advisor(s): Van hamme, H.), (Co-Advisor(s): Norouzi, Y.), 2015.
- Mirzaei, S., Norouzi, Y., Van hamme, H., "Two-stage blind audio source counting and separation of stereo instantaneous mixtures using Bayesian tensor factorisation", *IET Signal Processing*, vol. 9, no. 8, 2015, pp. 587-595.
- Mirzaei, S., Van hamme, H., Norouzi, Y., "Bayesian non-parametric matrix factorization for discovering words in spoken utterances", in *Proceedings WASPAA 2013 (20-23 October 2013, New Paltz, New York, USA)*, 2013, pp. 1-4.
- Mirzaei, S., Van hamme, H., Norouzi, Y., "Blind audio source counting and separation of anechoic mixtures using the multichannel complex NMF framework", *Signal Processing*, vol. 115, 2015, pp. 27-37.
- Mirzaei, S., Van hamme, H., Norouzi, Y., "Blind audio source separation of stereo mixtures using Bayesian non-negative matrix factorization", in *Proceedings EUSIPCO 2014 (1-5 September 2014, Lisbon, Portugal)*, 2014, pp. 1-5.
- Mirzaei, S., Van hamme, H., Norouzi, Y., "Blind speech source localization, counting and separation for 2-channel convolutive mixtures in a reverberant environment", in *Proceedings Interspeech 2014 (14-18 September 2014, Singapore)*, 2014, pp. 860-864.
- Mirzaei, S., Van hamme, H., Norouzi, Y., "Model order estimation using bayesian NMF for discovering phone patterns in spoken utterances", in *Proceedings Interspeech 2013 (25-29 August 2013, Lyon, France)*, 2013, pp. 1-5.
- Mirzaei, S., Van hamme, H., Norouzi, Y., "Under-determined reverberant audio source separation using Bayesian Non-negative Matrix Factorization", *Speech Communication*, vol. 81, 2016, pp. 129-137.
- Moens, M., "Argumentation mining: How can a machine acquire common sense and world knowledge?", *Argument & Computation*, vol. epub ahead of print, 2017.
- Moens, M., "Argumentation mining: How can a machine acquire world and common sense knowledge? (abstract of keynote lecture)", *Computational Models of Argument (6-12 September 2016, Potsdam, Germany)*, 2016, pp. 4.
- Moens, M., "Argumentation mining: Where are we now, where do we want to be and how do we get there?", in *Post-proceedings of the forum for information retrieval evaluation (FIRE 2013) (4-6 December 2013, New Delhi, India)*, 2013, art. nr. 2.

- Moens, M., "Information extraction from biomedical texts: Learning models with limited supervision", Proceedings of the sixth international workshop on health text mining and information analysis (Louhi 2015) (17 September 2015, Lisbon, Portugal), 2015, pp. 120.
- Moens, M., "Information extraction from text and beyond" in Cross-disciplinary advances in applied natural language processing: issues and approaches (Premier reference source), Boonthum-Denecke, C., McCarthy, P., Lamkin, T. Eds., Hershey, PA: IGI Global, 2012, pp. 24-39.
- Moens, M., "Learning representations for language understanding: experiences from the MUSE project", , 2015.
- Moens, M., "Machine understanding of language: how can a machine learn?", , 2016.
- Moens, M., Do Thi, N., Kordjamshidi, P., Ludwig, O., "Visualising the content of a text in a virtual world" in Kultur und Informatik: Cross Media, Busch, C., Sieck, J. Eds. Verlag Werner Hülsbusch, 2015, ISBN 978-3-86488-082-7, pp. 33-37.
- Moens, M., Li, J., Chua, T., "Conclusions and a roadmap for future developments" in Mining user generated content, Moens, M., Li, J., Chua, T. Eds., London, UK: CRC Press, 2014, ISBN 9781466557406, pp. 319-328.
- Moens, M., Li, J., Chua, T., "Mining user generated content and its applications" in Mining user generated content, Moens, M., Li, J., Chua, T. Eds., London, UK: CRC Press, 2014, ISBN 9781466557406, pp. 3-17.
- Moens, M., Li, J., Chua, T., Eds, "Mining user generated content", "Social media and social computing, London: Chapman and Hall/CRC, 2014. ISBN 9781466557406.
- Moens, M., Pastra, K., Saenko, K., Tuytelaars, T., "Vision and language integration meets multimedia fusion", In Proceedings of ACM Multimedia 2016 (15-19 October 2016, Amsterdam, The Netherlands), 2016, pp. 1493.
- Moens, M., Pastra, K., Saenko, K., Tuytelaars, T., Eds, "Proceedings of the 2016 ACM Workshop on Vision and Language Integration Meets Multimedia Fusion" ACM, 2016. ISBN 978-1-4503-4519-4.
- Moens, M., Viard-Gaudin, C., Zargayouna, H., Ramos-Terrades, O., Eds, "CORIA 2014 Actes de la onzième conférence en recherche d'information et applications" Association francophone de recherche d'information et applications, 2014. ISBN 978-2-37111-001-4.
- Moens, M., Vulic, I., "Monolingual and cross-lingual probabilistic topic models and their applications in information retrieval", in Lecture Notes in Computer Science (24-27 March 2013, Moscow, Russian Federation), 2013, pp. 875-878.
- Moens, M. & Vulic, I. "Monolingual and Cross-Lingual Probabilistic Topic Models and Their Applications in Information Retrieval". In Proceedings of the 35th European Conference on Information Retrieval (ECIR 2013).(Lecture Notes in Computer Science 7814) (pp. 775-778). Berlin: Springer 2013
- Montero Perez, M., "(W)ondertitels voor wie een nieuwe taal leert", iMinds nieuws, 2013.
- Montero Perez, M., "Enhancing vocabulary learning through captioned video: An eye-tracking study", (3 September 2013, Antwerpen), 2013.
- Montero Perez, M., "Eye-tracking and second language acquisition", (2 March 2015, Leuven), 2015.
- Montero Perez, M., "Eye-tracking as a means to gauge L2 learners' caption-reading behavior and attention allocation", (7-9 July 2014, 2014).
- Montero Perez, M., "Eye-Tracking as a Means to Gauge L2 Learners' Caption-Reading Behavior and Attention Allocation", in Research challenges in CALL (7-9 July 2014, Antwerpen), 2014, pp. 258-261.
- Montero Perez, M., "Incidental vocabulary learning from viewing: The role of frequency of occurrence, working memory, and vocabulary knowledge", (3-4 July 2017, Reading), 2017.
- Montero Perez, M., "Incidental vocabulary learning from viewing: the role of frequency of occurrence, working memory, and vocabulary size (accepted)", TBLT Conference (19-21 April 2017, Barcelona), 2017.
- Montero Perez, M., "Interview met Maribel Montero Perez: (W)ondertitels voor de vreemdetallearder?", Nieuwsbrief Taal en Onderwijs, 2014.
- Montero Perez, M., "Ondertiteling goed om taal te leren", De Morgen, 2013.
- Montero Perez, M., "Researching the effects of captioned video on L2 learning", (30 January 2015, Universitat de Barcelona), 2015.
- Montero Perez, M., "The effectiveness of different captioning techniques for foreign language listening comprehension and incidental vocabulary learning", (28-30 March 2012, Antwerp), 2012.
- Montero Perez, M., "The effectiveness of full captioning, keyword captioning, and full captioning with marked keywords for listening comprehension and incidental vocabulary learning", Subtitles and language learning (13-14 September 2012, Pavia), 2012, pp. 24-25.
- Montero Perez, M., "TV-kijken en woordenschat", (26 April 2017, 2017).
- Montero Perez, M., "TV-kijken en woordenschatverwerving", (21 April 2016, Gent), 2016.
- Montero Perez, M., "Vocabulary learning through captioned video: An eye-tracking study", (13 March 2014, VUB Brussel), 2013.

- Montero Perez, M., "Watch and Learn?! Five studies into the use and effectiveness of captioned video for L2 listening comprehension and vocabulary acquisition.", PhD Thesis (Advisor(s): Desmet, P.), (Co-Advisor(s): Peters, E.), 2013.
- Montero Perez, M., "Woordenschatverwerving door het bekijken van ondertitelde video: Een eye-trackingstudie", (3 May 2017, Universiteit Gent), 2017.
- Montero Perez, M., Desmet, P., "Enjeux et défis de la création d'un environnement d'apprentissage mobile: le cas de "MAPLE""", *Romanesque*, vol. 37, no. 2, 2012, pp. 3-13.
- Montero Perez, M., Desmet, P., "Exploring the potential of keyword captions, full captions, and full captions with marked keywords for listening comprehension and incidental vocabulary", *CALL: Using, learning, knowing* (22-25 August 2012, Gothenburg), 2012, pp. 189-191.
- Montero Perez, M., Desmet, P., "The effect of input enhancement in L2 listening on incidental vocabulary learning: A review", in *Procedia - Social and Behavioral Sciences* (8-11 September 2010, Bordeaux), Macaire, D., Boulton, A. Eds., 2012, pp. 153-157.
- Montero Perez, M., Desmet, P., "The role of input for foreign language learning: From theory to pedagogical applications", (27 April 2017, Leuven), 2017.
- Montero Perez, M., Paulussen, H., Macken, L., Desmet, P., "From input to output: The potential of parallel corpora for CALL", *Language Resources and Evaluation*, vol. 48, no. 1, 2014, pp. 165-189.
- Montero Perez, M., Peters, E., "Methodological choices in studies into vocabulary acquisition through audio(-visual) input: a systematic review", (7-8 July 2016, Nottingham, UK), 2016.
- Montero Perez, M., Peters, E., Clarebout, G., Desmet, P., "Effects of captioning on video comprehension and incidental vocabulary learning", *Language Learning & Technology*, vol. 18, no. 1, 2014, pp. 118-141.
- Montero Perez, M., Peters, E., Desmet, P., "Enhancing vocabulary learning through captioned video: An eye-tracking study", *The Modern Language Journal*, vol. 99, no. 2, 2015, pp. 308-328.
- Montero Perez, M., Peters, E., Desmet, P., "Incidental and intentional vocabulary learning through video with full or (glossed) keyword captions", (22-25 March 2014, Portland, US), 2014.
- Montero Perez, M., Peters, E., Desmet, P., "Incidental and intentional vocabulary learning through video with glossed captions", *EUROSLA 24* (3-6 September 2014, York), 2014, pp. 114-114.
- Montero Perez, M., Peters, E., Desmet, P., "Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension", *ReCALL: Journal of EuroCALL*, vol. 26, no. 1, 2014, pp. 21-43.
- Montero Perez, M., Peters, E., Desmet, P., "Vocabulary test announcement and learners' use of (keyword) captioned video: An eye-tracking study", *Eurosla 23: Book of Abstracts* (28-31 August 2013, Amsterdam, The Netherlands), 2013, pp. 124-125.
- Montero Perez, M., Van Den Noortgate, W., Desmet, P., "Captioned video for L2 listening and vocabulary learning: A meta-analysis", *System*, vol. 41, no. 3, 2013, pp. 720-739.
- Moser, C., Tscheligi, M., Zaman, B., Vanden Abeele, V., Geurts, L., Vandewaetere, M., Markopoulos, P., Wyeth, P., Hofstätter, J., "Let's talk about failures: why was the game for children not a success?", *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (27 April - 2 May, Paris), 2013, pp. 3199-3202.
- Moser, M., Tscheligi, C., Zaman, B., Vanden Abeele, V., Geurts, L., Vandewaetere, M., Markopoulos, P., Wyeth, P., Eds, "Special Issue: Learning from Failures in Game Design for Children", *International Journal of Child - Computer Interaction*, vol. 2,2" Elsevier Inc., 2014.
- Murillo Mange, Y., Van den Bergh, B., Beysens, J., Bertrand, A., Dehaene, W., Patrinos, P., Tuytelaars, T., Sabariego, R., Verhelst, M., Wambacq, P., Pollin, S., "Multidisciplinary Learning through Implementation of the DVB-S2 Standard", *IEEE Communications Magazine*, vol. 55, no. 5, 2017, pp. 124-130.
- Noreillie, A., Chainaye, L., Desmet, P., "The online language learning platform FRANEL: Effective use and perceived added value", (22-25 August 2012, Gothenburg, Sweden), 2012.
- Noreillie, A., Grisez, V., Desmet, P., "(Semi)authentic audio-visual materials for the A2 level in the online language learning environment FRANEL: Pitfalls and challenges", in *Procedia - Social and Behavioral Sciences* (8-11 September 2010, Bordeaux), Macaire, D., Boulton, A. Eds., 2012, pp. 164-168.
- Noreillie, A., Kestemont, B., Heylen, K., Desmet, P., Peters, E., "An empirical investigation of the relationship between L2 vocabulary size and listening comprehension at the B1 CEFR-level in French and English", (30 May 2016 - 1 June 2016, Calgary, Canada), 2016.
- Noreillie, A., Kestemont, B., Peters, E., Desmet, P., Heylen, K., "VALILEX - Theoretical and empirical validation of lexical competence in English and French within the Common European Framework of Reference", (6 March 2015, Nijmegen, The Netherlands), 2015.
- Nurani Venkitasubramanian, A., "Weakly Supervised Machine Learning Algorithms for Object Recognition in the Wild and Entity Linking in Videos", PhD Thesis (Advisor(s): Moens, M., Tuytelaars, T.), 2017.

- Nurani Venkitasubramanian, A., Moens, M., "Estimating the breadth of search queries", in Proceedings of the 10th International RIAO conference on open research areas in information retrieval (22-24 May 2013, Lisbon, Portugal), 2013, pp. 109-112.
- Nurani Venkitasubramanian, A., Moens, M., "Selection of search facets", in Proceedings of CORIA 2014 (3-5 April 2013, Neuchâtel, Switzerland), 2013, pp. 73-84.
- Nurani Venkitasubramanian, A., Moens, M., "Summarization and expansion of search facets", Proceedings of the thirteenth Dutch-Belgian information retrieval workshop (26 April 2013, Delft, The Netherlands), 2013, pp. 50-51.
- Nurani Venkitasubramanian, A., Tuytelaars, T., Moens, M., "Entity linking across vision and language", *Multimedia Tools and Applications*, vol. 76, no. 21, 2017, pp. 22599-22622.
- Nurani Venkitasubramanian, A., Tuytelaars, T., Moens, M., "Learning to recognize animals by watching documentaries: using subtitles as weak supervision", in Proceedings of the 6th Workshop on Vision and Language (VL'17) at EACL 2016 (4 April 2017, Valencia), 2017, pp. 21-30.
- Nurani Venkitasubramanian, A., Tuytelaars, T., Moens, M., "Wildlife recognition in nature documentaries with weak supervision from subtitles and external data", *Pattern Recognition Letters*, vol. 81, 2016, pp. 63-70.
- Onrust, L., van den Bosch, A., Van hamme, H., "Improving cross-domain n-gram language modelling with skipgrams", in Proceedings ACL 2016 (7-12 August 2016, Berlin, Germany), 2016, pp. 137-142.
- Ons, B., "The Self-taught Speech Interface", PhD Thesis (Advisor(s): Van hamme, H.), (Co-Advisor(s): Gemmeke, J.), 2015.
- Ons, B., Gemmeke, J., Van hamme, H., "An adaptive vocal user-interface learning from speech examples of the user", (11 September 2013, Heverlee, Belgium), 2013.
- Ons, B., Gemmeke, J., Van hamme, H., "Fast vocabulary acquisition in an NMF-based self-learning vocal user interface", *Computer Speech and Language*, vol. 28, no. 4, 2014, pp. 997-1017.
- Ons, B., Gemmeke, J., Van hamme, H., "Label noise robustness and learning speed in a self-learning vocal user interface", in Proceedings IWSDS 2012 (28-30 November 2012, Paris, France), 2012, pp. 219-229.
- Ons, B., Gemmeke, J., Van hamme, H., "NMF-based keyword learning from scarce data", in Proceedings ASRU 2013 (8-12 December 2013, Olomouc, Czech Republic), 2013, pp. 392-397.
- Ons, B., Gemmeke, J., Van hamme, H., "The self-taught vocal interface", *EURASIP Journal on Audio, Speech and Music Processing*, vol. 43, 2014, pp. 1-16.
- Ons, B., Tessema, N., van de Loo, J., Gemmeke, J., De Pauw, G., Daelemans, W., Van hamme, H., "A self learning vocal interface for speech-impaired users", in Proceedings SLPAT 2013 (21-22 August 2013, Grenoble, France), 2013, pp. 73-81.
- Oostdijk, N., Reynaert, M., Hoste, V., Schuurman, I., "The Construction of a 500-million-word Reference Corpus of Contemporary Written Dutch" in *Essential Speech and Language Technology for Dutch: resources, tools and applications*, Spyns, P., Odijk, J. Eds. Springer, 2013, ISBN 978-3-642-30909-0, pp. 219-247.
- Paiva Proenca, K., Demuyndck, K., Van Compennolle, D., "Automatic syllabification using segmental conditional random fields", *Computational Linguistics in the Netherlands Journal*, vol. 3, 2013, pp. 34-48.
- Paiva Proenca, K., Demuyndck, K., Van Compennolle, D., "Designing syllable models for an HMM based speech recognition system", in *Lecture notes in artificial intelligence* (23-27 August 2016, Budapest, Hungary), 2016, pp. 216-223.
- Paulussen, H., Anné, Y., Gouder, E., Suys, V., Lauwers, J., Driesen, J., Cornillie, F., "Building a generic morphological tagger for Latin", (16-20 May 2017, Flagstaff, Arizona), 2017.
- Paulussen, H., Bonachela Capdevila, F., Debevere, P., Montero Perez, M., Vanbrabant, M., De Neve, W., De Wannemacker, S., "Automatic document enrichment for language learners", (11-12 November 2013, Stellenbosch), 2013.
- Paulussen, H., Debevere, P., Bonachela Capdevila, F., Montero Perez, M., Vanbrabant, M., De Neve, W., De Wannemacker, S., "Building an NLP pipeline within a digital publishing workflow", (17 January 2014, 2014).
- Paulussen, H., Debevere, P., Bonachela Capdevila, F., Montero Perez, M., Vanbrabant, M., De Neve, W., De Wannemacker, S., "Building an NLP pipeline within a digital publishing workflow", *Computational Linguistics in the Netherlands Journal*, vol. 4, 2014, pp. 71-84.
- Paulussen, H., Macken, L., Vandeweghe, W., Desmet, P., "Dutch Parallel Corpus: A Balanced Parallel Corpus for Dutch-English and Dutch-French" in *Essential Speech and Language Technology for Dutch.*, Spyns, P., Odijk, J. Eds., Heidelberg, New York, Dordrecht, London: Springer Verlag, 2013, ISBN 978-3-642-30909-0, pp. 185-199.
- Paulussen, H., Vanbrabant, M., Haesendonck, G., "From paper dictionary to electronic knowledge base", *CLIN proceedings* (5-6 February 2015, Antwerp), 2015.
- Pelemans, J., "Efficient Language Modeling for Automatic Speech Recognition", PhD Thesis (Advisor(s): Wambacq, P., Van hamme, H.), 2017.

- Pelemans, J., De Laet, B., Van hamme, H., Wambacq, P., "The effect of word similarity on N-gram language models in northern and southern Dutch", Abstract book CLIN 2014 (17 January 2014, Leiden, The Netherlands), 2014, pp. 84.
- Pelemans, J., Demuynck, K., Van hamme, H., Wambacq, P., "Coping with language data sparsity: semantic head mapping for compound words", in Proceedings ICASSP 2014 (4-9 May 2014, Florence, Italy), 2014, pp. 141-145.
- Pelemans, J., Demuynck, K., Van hamme, H., Wambacq, P., "Improving N-gram probability estimates by compound-head clustering", in Proceedings ICASSP 2015 (19-24 April 2015, Brisbane, Australia), 2015, pp. 5221-5225.
- Pelemans, J., Demuynck, K., Van hamme, H., Wambacq, P., "Speech recognition web services for Dutch", in Proceedings LREC 2014 (26-31 May 2014, Reykjavik, Iceland), 2014, pp. 3041-3044.
- Pelemans, J., Demuynck, K., Van hamme, H., Wambacq, P., "The effect of word similarity on N-gram language models in Northern and Southern Dutch", Computational Linguistics in the Netherlands Journal, vol. 4, 2014, pp. 91-104.
- Pelemans, J., Demuynck, K., Wambacq, P., "A layered approach for Dutch large vocabulary continuous speech recognition", in Proceedings ICASSP'2012 (25-30 March 2012, Kyoto, Japan), 2012, pp. 4421-4424.
- Pelemans, J., Demuynck, K., Wambacq, P., "Dutch automatic speech recognition on the web: Towards a general purpose system", in Proceedings Interspeech 2012 (9-13 September 2012 (Best Show & Tell Presentation), Portland, Oregon, USA), 2012, art. nr. 453.
- Pelemans, J., Shazeer, N., Chelba, C., "Pruning sparse non-negative matrix N-gram language models", in Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany), 2015, pp. 1433-1437.
- Pelemans, J., Shazeer, N., Chelba, C., "Sparse non-negative matrix language modeling", (1-5 November 2016, Austin, Texas, USA), 2016.
- Pelemans, J., Shazeer, N., Chelba, C., "Sparse non-negative matrix language modeling", Transactions of the Association for Computational Linguistics, vol. 4, 2016, pp. 329-342.
- Pelemans, J., Van hamme, H., Wambacq, P., "Translation-based word clustering for language models", CLIN 2015 book of abstracts (5-6 February 2015, Antwerp, Belgium), 2015, pp. 60.
- Pelemans, J., Vanallemeersch, T., Demuynck, K., Van hamme, H., Wambacq, P., "Efficient language model adaptation for automatic speech recognition of spoken translations", in Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany), 2015, pp. 2262-2266.
- Pelemans, J., Vanallemeersch, T., Demuynck, K., Verwimp, L., Van hamme, H., Wambacq, P., "Language model adaptation for ASR of spoken translations using phrase-based translation models and named entity models", in Proceedings ICASSP 2016 (20-25 March 2016, Shanghai, China), 2016, pp. 5985-5989.
- Pelemans, J., Vanallemeersch, T., Demuynck, K., Verwimp, L., Van hamme, H., Wambacq, P., "Machine translation-based language model adaptation for automatic speech recognition of spoken translations", Abstract book CLIN26 (18 December 2015, Amsterdam, The Netherlands), 2015, pp. 26.
- Pelemans, J., Verwimp, L., Demuynck, K., Van hamme, H., Wambacq, P., "SCALE: A scalable language engineering toolkit", in Proceedings LREC 2016 (23-28 May 2016, Portorož, Slovenia), 2016, pp. 3868-3871.
- Peters, E., Kestemont, B., Noreillie, A., Desmet, P., Heylen, K., "The effect of naturalistic language learning on learners' English and French vocabulary size and attitudes", (16-18 September 2015, Leuven, Belgium), 2015.
- Peters, E., Montero Perez, M., "Vocabulary acquisition through audio(-visual) input: how to measure learning gains?", (11-12 June 2015, Barcelona, Spain), 2015.
- Peters, S., Clarebout, G., Aertgeerts, B., Roex, A., "Offering classroom-based learning materials via online technology during internship: A qualitative study", (24-26 August 2016, Dijon, France), 2016.
- Peters, S., Clarebout, G., Aertgeerts, B., Roex, A., "Supporting students during their internship via an Electronic Performance Support System", (24-26 August 2016, Dijon, France), 2016.
- Pham The, P., Deschacht, K., Tuytelaars, T., Moens, M., "Naming persons in video: using the weak supervision of textual stories", Journal of Visual Communication and Image Representation, vol. 24, no. 7, 2013, pp. 944-955.
- Pham, P., "Algorithms for Cross-Media Alignment of Equivalent Content (Algoritmen voor cross-mediale alignering van equivalente inhoud)", PhD Thesis (Advisor(s): Moens, M., Tuytelaars, T.), 2012. ISBN 978-94-6018-576-2.
- Poelmans, S., Truyen, F., Stockman, C., "ICT Skills and Computer Self-Efficacy of Higher Education Students", in INTED2012 Proceedings 6th International Technology, Education and Development Conference (5-7th March 2012, Valencia, Spain), Gómez Chova, L., López Martínez, A., Candel Torres, I. Eds., 2012, art. nr. 1303, pp. 1123-1132.
- Poels, Y., Annema, J., Zaman, B., Cornillie, F., "Developing user-centered concepts for language learning video games", in Proceedings of The Fifth International Conference on Advances in Computer-Human Interactions ACHI 2012 (30/01-04/02/11, Valencia, Spain), 2012.

- Poorjam, A., Bahari, M., Van hamme, H., "Multitask speaker profiling for estimating age, height, weight and smoking habits from spontaneous telephone speech signals", in Proceedings ICCCKE 2014 (29-30 October 2014, Mashhad, Iran), 2014, pp. 7-12.
- Poorjam, A., Bahari, M., Van hamme, H., "Speaker weight estimation from speech signals using a fusion of the i-vector and NFA frameworks", in Proceedings AISP 2015 (3-5 March 2015, Mashhad, Iran), 2015, pp. 1-6.
- Poorjam, A., Bahari, M., Vasilakakis, V., Van hamme, H., "Height estimation from speech signals using i-vectors and least-squares support vector regression", in Proceedings TSP 2014 (1-3 July 2014, Berlin, Germany), 2014, art. nr. 201, pp. 1-5.
- Poulisse, G., "Complex Semantic Concept Detection in Video (Complex semantische concept herkenning in video)", PhD Thesis (Advisor(s): Moens, M., Van Gool, L.), 2012.
- Poulisse, G., Patsis, Y., Moens, M., "Unsupervised scene detection and commentator building using multi-modal chains", *Multimedia Tools and Applications*, vol. 70, no. 1, 2014, pp. 159-175.
- Provoost, T., Moens, M., "Detecting relations in the gene regulation network", in Proceedings of BioNLP shared task 2013 workshop: the Genia event extraction shared task (8-9 August 2013, Sofia, Bulgaria), 2013, pp. 135-138.
- Provoost, T., Moens, M., "Semi-supervised learning for the BioNLP gene regulation network", *BMC Bioinformatics*, vol. 16 (Suppl 10), 2015, art.nr. S4.
- Pustejovsky, J., Kordjamshidi, P., Moens, M., Levine, A., Dworman, S., Yocum, Z., "SemEval-2015 task 8: SpaceEval", in Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015) (4-5 June 2015, Denver, Colorado), 2015, pp. 884-894.
- Quan, L., Kolomiyets, O., Moens, M., "KU Leuven at HOO-2012: A hybrid approach to detection and correction of determiner and preposition errors in non-native English text", in Proceedings of the NAACL-HLT 2012 workshop on innovative use of NLP for building educational applications (7 June 2012, Montreal, Canada), 2012, pp. 263-271.
- Raes, A., "Op onderzoek met WISE in de klas. Workshop", (15 november 2014, Kortrijk, België), 2014.
- Raes, A., "The impact of "computer-supported collaborative inquiry" for science learning in secondary education", PhD Thesis (Advisor(s): Schellens, T.), 2015.
- Raes, A., "WISE in de klas. Workshop", (10 maart 2012, Oostende, België), 2012.
- Raes, A., Adams, B., Montrieux, H., Schellens, T., "Using Twitter in higher education: What are students' initial perceptions and experiences?", in Proceedings of the 3rd European Conference on Social Media (July 12-13, 2016, Caen: France), 2016, pp. 324-331.
- Raes, A., Ameloot, E., Schellens, T., "Synergistic Scaffolding in Technology-Enhanced Learning: How learning analytics can support teaching and learning", (11-13 October 2017, Leuven, Belgium), 2017.
- Raes, A., Depaepe, F., "Smart Education" Hoe kan technologie kwaliteitsvol onderwijs ondersteunen?", , 2016.
- Raes, A., Desmet, P., Depaepe, F., "Technology-enhanced collaborative learning for improved interactivity, collaboration, and flexibility in higher education and corporate training", *Making a difference: Prioritizing equity and access*. (17-22 June 2017, Philadelphia (PA), USA), 2017, pp. 942-943.
- Raes, A., Montrieux, H., Schellens, T., Vangestel, S., De Clercq, P., "Weblectures voor meer studiesucces", (23 mei 2013, Gent, België), 2013.
- Raes, A., Schellens, T., "Can we increase students' motivation to learn science by means of web-based collaborative inquiry?", *To see the world and a grain of sand: Learning across levels of space, time, and scale: CSCL 2013 Conference Proceedings* (15-19 Juni 2013, Wisconsin, Madison, USA), 2013, pp. 335-336.
- Raes, A., Schellens, T., "In search of a synergy between multilevel scaffolding of technology-enhanced inquiry in complex classrooms", (27 January - 1 February 2013, Villard-de-Lans, France), 2013.
- Raes, A., Schellens, T., "In search of a synergy between multilevel scaffolding of web-based inquiry in complex classrooms. Symposium conducted at the 15th biennial EARLI conference for research on learning and instruction", *EARLI 2013, Book of Abstracts* (2013, Munich, Germany), 2013, pp. 230-231.
- Raes, A., Schellens, T., "Teacher-led class interventions can make the difference in computer-supported collaborative science learning", *AERA Online Paper Repository* (15-21 April, 2015, Chicago, USA), 2015.
- Raes, A., Schellens, T., "The effects of teacher-led class interventions during technology-enhanced science inquiry on students' knowledge achievement and basic need satisfaction", (18-20 August 2014, Malmö, Sweden), 2014.
- Raes, A., Schellens, T., "The effects of teacher-led class interventions during technology-enhanced science inquiry on students' knowledge integration and basic need satisfaction", *Computers and Education*, vol. 92-93, 2016, pp. 125-141.

- Raes, A., Schellens, T., "The impact of web-based inquiry in secondary science education on students' motivation for science learning", in *Procedia: Social and Behavioral Sciences* (OCT 10-13, 2012, Istanbul: TURKEY), 2012, pp. 1332-1339.
- Raes, A., Schellens, T., "Unraveling the motivational effects and challenges of web-based collaborative inquiry learning across different groups of learners", *Educational Technology Research and Development*, vol. 63, no. 3, 2015, pp. 405-430.
- Raes, A., Schellens, T., De Wever, B., "Scripting collaboration in web-based collaborative inquiry in face-to-face classroom settings", *EARLI SIG 20 Computer-Supported Inquiry Learning. Conference Summaries*, Bochum, Germany), 2012, pp. 26-27.
- Raes, A., Schellens, T., De Wever, B., "Web-based collaborative inquiry to bridge gaps in secondary science education", *Journal of the Learning Sciences*, vol. 23, no. 3, 2014, pp. 316-347.
- Raes, A., Schellens, T., De Wever, B., Benoit, D., "Promoting metacognitive regulation through collaborative problem solving on the web: When scripting does not work", *Computers in Human Behavior*, vol. 58, 2016, pp. 325-342.
- Raes, A., Schellens, T., De Wever, B., Kollar, I., Wecker, C., Fischer, F., Tissenbaum, M., Slotta, J., Peters, V., Songer, N., "Scripting science inquiry learning in CSCL classrooms. Symposium conducted at the 10th International Conference of the Learning Sciences (ICLS 2012)", in *Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012) (2-6 July, 2012, Sydney, Australia)*, van Aalst, J., Thompson, K., Jacobson, M., Reimann, P. Eds., 2012, pp. 118-125.
- Raes, A., Schellens, T., De Wever, B., Vanderhoven, E., "Scaffolding information problem solving in web-based collaborative inquiry learning", *Computers and Education*, vol. 59, no. 1, 2012, pp. 82-94.
- Raes, A., Vanderhoven, E., "Waarom en hoe technologie inzetten als leermiddel voor de student van vandaag?", (3 December 2015, Kortrijk, België), 2015.
- Raes, A., Vanderhoven, E., Schellens, T., "Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education", *Studies in Higher Education*, vol. 40, no. 1, 2015, pp. 178-193.
- Raes, A., Vens, C., Vanneste, P., & Depaepe, F.(2018). Teaching for versus through Problem Solving: Impact on Teaching and Learning. *International Conference of the Learning Sciences (ICLS) 2018; 13; London, UK.*
- Raes, A., Wichmann, A., Ludvigsen, S., Stromme, T., Kollar, I., Schellens, T., Slotta, J., Linn, M., "WISE in Europe: Discovering the Web-based Inquiry Science Environment from three perspectives", *EARLI 2015, Book of Abstracts (25-29 August 2015, Limassol, Cyprus)*, 2015, pp. 617-617.
- Renkens, V., Janssens, S., Ons, B., Gemmeke, J., Van hamme, H., "Acquisition of ordinal words using weakly supervised NMF", in *Proceedings SLT 2014 (7-10 December 2014, South Lake Tahoe, Nevada, USA)*, 2014, pp. 1-6.
- Renkens, V., Tomar, V., Van hamme, H., "Incrementally learn the relevance of words in a dictionary for spoken language acquisition", in *Proceedings SLT 2016 (13-16 December 2016, San Diego, California, USA)*, 2016, pp. 144-150.
- Renkens, V., Van hamme, H., "Automatic relevance determination for nonnegative dictionary learning in the Gamma-Poisson model", *Signal Processing*, vol. 132, 2017, pp. 121-133.
- Renkens, V., Van hamme, H., "Mutually exclusive grounding for weakly supervised Non-negative Matrix Factorisation", in *Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany)*, 2015, pp. 1388-1392.
- Renkens, V., Van hamme, H., "Weakly supervised learning of Hidden Markov Models for spoken language acquisition", *IEEE-acm Transactions on Audio Speech and Language Processing*, vol. 25, no. 2, 2017, pp. 285-295.
- Reynaert, M., Schuurman, I., Hoste, V., Oostdijk, N., Van Gompel, M., "Beyond SoNaR: towards the facilitation of large corpus building efforts", in *Proceedings of the Eighth International conference on Language Resources and Evaluation (LREC, Istanbul (Turkey))*, 2012, pp. 2897-2904.
- Roose, H., Van Nieuwenhuyse, K., Wils, K., Depaepe, F., Verschaffel, L., Frère, G., Loquet, O., Smits, W., Vandevanne, K., Van Hooydonck, I., Verlinden, E., "Over een reflectieve omgang met bronnen in het onderwijs in geschiedenis en cultuurwetenschappen", (26-27 March 2015, Arnhem), 2015.
- Roose, H., Van Nieuwenhuyse, K., Wils, K., Depaepe, F., Verschaffel, L., Frère, G., Loquet, O., Smits, W., Vandevanne, K., Van Hooydonck, I., Verlinden, E., "Over een reflectieve omgang met bronnen in het onderwijs in geschiedenis en cultuurwetenschappen", (26 February 2015, Hasselt), 2015.
- Rotsaert, T., Raes, A., Schellens, T., "Formative peer assessment/feedback using anonymous mobile response technology: Exploring interpersonal variables and teacher guidance", *AERA Online Paper Repository (15-21 April, 2015, Chicago, USA)*, 2015.

- Ruette T., Ehret K., Szmrecsanyi B. (2016). A lectometric analysis of aggregated lexical variation in written Standard English with Semantic Vector Space models.. *International Journal of Corpus Linguistics*, 21 (1), 48-79.
- Ruiter, M., Beijer, L., Cucchiari, C., Krahmer, E., Rietveld, T., Strik, H., Van hamme, H., "Human language technology and communicative disabilities: requirements and possibilities for the future", *Language Resources and Evaluation*, vol. 46, no. 1, 2012, pp. 143-151.
- Saggion, H., Ferrés, D., Sevens, L., Schuurman, I., Ripollés, M., Rodríguez, O., "Able to Read My Mail: An Accessible e-Mail Client with Assistive Technology", in *Proceedings of the 14th Web For All (W4A) Conference – The Future of Accessible Work (2-4 April 2017, Perth)*, Bailey, C., Gkatzidou, V. Eds., 2017, art. nr. 5, pp. 1-4.
- Sahraeian, R., "Acoustic Modeling of Under-resourced Languages", PhD Thesis (Advisor(s): Van Compernelle, D.), 2017.
- Sahraeian, R., Van Compernelle, D., "A study of rank-constrained multilingual DNNS for low-resource ASR", in *Proceedings ICASSP 2016 (20-25 March 2016, Shanghai, China)*, 2016, pp. 5420-5424.
- Sahraeian, R., Van Compernelle, D., "A study of supervised intrinsic spectral analysis for TIMIT phone classification", in *Proceedings ASRU 2013 (8-12 December 2013, Olomouc, Czech Republic)*, 2013, pp. 256-260.
- Sahraeian, R., Van Compernelle, D., "Exploiting sequential low-rank factorization for multilingual DNNS", in *Proceedings ICASSP 2017 (5-9 March 2017, New Orleans, USA)*, 2017, pp. 5025-5029.
- Sahraeian, R., Van Compernelle, D., "Using weighted model averaging in distributed multilingual DNNS to improve low resource ASR", in *Procedia Computer Science (9-12 May 2016, Yogyakarta, Indonesia)*, 2016, pp. 152-158.
- Sahraeian, R., Van Compernelle, D., de Wet, F., "On Using Intrinsic Spectral Analysis for Low-resource Languages", in *Proceedings SLTU 2014 (14-16 May 2014, Saint-Petersburg, Russia)*, 2014, pp. 61-65.
- Sahraeian, R., Van Compernelle, D., de Wet, F., "Under-resourced speech recognition based on the speech manifold", in *Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany)*, Sahraeian Eds., 2015, pp. 1255-1259.
- Sahraeian, R., Van Compernelle, D., de Wet, F., "Using generalized maxout networks and phoneme mapping for low resource ASR– a case study on Flemish-Afrikaans", in *Proceedings PRASA-RobMech (26-27 November 2015, Port Elizabeth, South Africa)*, 2015, pp. 112-117.
- Sainath, T., Ramabhadran, B., Nahamoo, D., Kanevsky, D., Van Compernelle, D., Demuynck, K., Gemmeke, J., Bellegarda, J., Sundaram, S., "Exemplar-based processing for speech recognition: An overview", *IEEE Signal Processing Magazine*, vol. 29, no. 6, 2012, pp. 98-113.
- Šarić, F., Dalbelo Bašić, B., Moens, M., Šnajder, J., "Multi-label classification of Croatian legal documents using EuroVoc thesaurus", in *Proceedings of SPlET - Semantic processing of legal texts: Legal resources and access to law workshop (27 May 2014, Reykjavik, Iceland)*, 2014, art. nr. <https://sites.google.com/site/splet2014workshop/workshop-programme>.
- Sauro, S., Cornillie, F., Buendgens-Kosten, J., "FanTALES: A needs analysis for multilingual digital storytelling tasks in 21st century European classrooms", (16-20 May 2017, Northern Arizona University, Flagstaff, AZ), 2017.
- Schuurman, I., Vandeghinste, V., Saggion, H. Eds., "Proceedings of the LREC workshop "Improving Social Inclusion using NLP: Tools and resources", " (23 May 2016, Portorož, Slovenia), ,2016, pp. 1-38
- Schuurman, I., Windhouwer, M., Ohren, O., Daniel, Z., "CLARIN Concept Registry: The New Semantic Registry", in *Selected Papers from the CLARIN Annual Conference 2015, October 14–16, 2015, Wroclaw, Poland (15-17 October 2015, Wroclaw, Poland)*, 2016, art. nr. 005, pp. 62-70.
- Segers, W., Kockaert, H., Wylín, B., "Evaluation de la traduction complète et évaluation sur la base d'items sélectionnés", (19-20 novembre 2015, Bruxelles/Louvain-la-Neuve, Belgique), 2015.
- Sevens, L., Daems, J., De Vlieghe, A., Schuurman, I., Vandeghinste, V., Van Eynde, F., "Building an Accessible Pictograph Interface for Users with Intellectual Disabilities" in *Harnessing the Power of Technology to Improve Lives*, Cudd, P., de Witte, L. Eds. IOS Press, 2017, ISBN 1614997985, pp. 870-877.
- Sevens, L., Jacobs, G., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Improving Text-to-Pictograph Translation Through Word Sense Disambiguation", in *Proceedings of the 5th Joint Conference on Lexical and Computational Semantics (11-12 August 2016, Berlin)*, 2016, pp. 1-5.
- Sevens, L., O'Flaherty, J., Schuurman, I., Vandeghinste, V., Van Eynde, F., "E-Inclusion of Functionally Illiterate Users by the use of Language Technology", in *Proceedings of the 2nd Conference on Engineering4Society: Raising awareness for the societal role of engineering (15-16 september 2016, Heverlee)*, 2016, pp. 91-95.
- Sevens, L., Schuurman, I., De Vlieghe, A., Daems, J., "Unity is Strength: How User Feedback Influences Technical Decisions in the Able to Include Project", *Proceedings of the NNDR 14th Research Conference (3-5 May 2017, Orebro)*, 2017.

- Sevens, L., Schuurman, I., O'Flaherty, J., Vandeghinste, V., Van Eynde, F., "Pictographs to the Rescue! Social Media for Functionally Illiterate Users", Proceedings of the 2016 Lancaster Disability Studies Conference (6-8 September 2016, Lancaster), 2016.
- Sevens, L., Schuurman, I., Vandeghinste, V., Van Eynde, F., "Automatic Translation with Pictographs to Serve People with IDD", Proceedings of the NNDR 13th Research Conference (8 May 2015, Bryggen), 2015.
- Sevens, L., Schuurman, I., Vandeghinste, V., Van Eynde, F., "Pictograph Translation Technologies for Improving E-Inclusion", Book of abstracts PLIN 2016 (12 May 2016, Louvain-la-Neuve), 2016.
- Sevens, L., Vanallemeersch, T., Schuurman, I., Vandeghinste, V., Van Eynde, F., "Automated Spelling Correction for Dutch Internet Users with Intellectual Disabilities", in Proceedings of 1st Workshop on Improving Social Inclusion using NLP: Tools and Resources (23 May 2016, Portorož, Slovenia), Schuurman, I., Vandeghinste, V., Saggion, H. Eds., 2016, pp. 11-19.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Extending a Dutch Text-to-Pictograph Converter to English and Spanish", in Proceedings of 6th Workshop on Speech and Language Processing for Assistive Technologies (11 September 2015, Dresden, Germany), 2015, pp. 1-8.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Make it Simple, But Significant: Improved Text-to-Pictograph Translation for People with Intellectual Disabilities", Proceedings of the NNDR 14th Research Conference (3-5 May 2017, Orebro), 2017.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Natural Language Generation from Pictographs", Book of abstracts CLIN26 (18 December 2015, Amsterdam), 2015.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Natural Language Generation from Pictographs", in Proceedings of the 15th European Workshop on Natural Language Generation (ENLG) (10-11 September 2015, Brighton, UK), Belz, A., Gatt, A., Portet, F., Purver, M. Eds., 2015, pp. 71-75.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Simplified Text-to-Pictograph Translation for People with Intellectual Disabilities" in Natural Language Processing and Information Systems (NLDB 2017), Frasincar, F., Ittoo, A., Nguyen, L., Métais, E. Eds., Cham: Springer International Publishing, 2017, ISBN 978-3-319-59569-6, pp. 185-196.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Syntactic Simplification for Improved Text-to-Pictograph Translation", Book of abstracts CLIN27 (10 February 2017, Leuven), 2017.
- Sevens, L., Vandeghinste, V., Schuurman, I., Van Eynde, F., "Text-To-Pictograph Translation for Six Language Pairs", Book of abstracts CLIN25 (6 February 2015, Antwerp), 2015.
- Sevens, L., Vandeghinste, V., Van Eynde, F., "Improving the Precision of Synset Links Between Cornetto and Princeton WordNet", in Proceedings of Workshop on Lexical and Grammatical Resources for Language Processing (24 August 2014, Dublin), 2014, pp. 120-126.
- Shazeer, N., Pelemans, J., Chelba, C., "Sparse non-negative matrix language modeling For skip-grams", in Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany), 2015, pp. 1428-1432.
- Sheng, S., Moens, M., "Simple baseline models for multimodal question answering in the cultural heritage domain" in Culture and Computer Science, Busch, C., Kussung, C., Sieck, J. Eds. Verlag Werner Hülsbusch, 2017, ISBN 978-3-86488-119-0, pp. 119-132.
- Sheng, S., Van Gool, L., Moens, M., "A dataset for multimodal question answering in the cultural heritage domain", in Proceedings of the COLING 2016 Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH) (11 December 2016, Osaka, Japan), 2016, pp. 10-17.
- Shi, Y., Larson, M., Pelemans, J., Jonker, C., Wambacq, P., Wiggers, P., Demuynck, K., "Integrating meta-information into recurrent neural network language models", Speech Communication, vol. 73, 2015, pp. 64-80.
- Shrestha, N., Moens, M., "Semi-automatic alignment of predicates between speech and OntoNotes data", in Proceedings of the 10th Language Resources and Evaluation Conference (23-28 May 2016, Portorož, Slovenia), 2016, pp. 1397-1401.
- Shrestha, N., Nurani Venkitasubramanian, A., Moens, M., "Key event detection in video using ASR and visual data", in Proceedings of the COLING Workshop on Vision and Language (VL'14) (23 August 2014, Dublin, Ireland), 2014, pp. 46-53.
- Shrestha, N., Vulic, I., Moens, M., "An IR-inspired approach to recovering named entity tags in broadcast news", (11 September 2013, Leuven), 2013, abstract nr.
<http://www.kuleuven.be/LICT/index.php/activities/details/86-lict-scientific-symposium-on-adaptivity-in-ict>.
- Shrestha, N., Vulic, I., Moens, M., "An IR-inspired approach to recovering named entity tags in broadcast news", in Lecture Notes in Computer Science (7-9 October 2013, Limassol, Cyprus), 2013, pp. 45-57.

- Shrestha, N., Vulic, I., Moens, M., "Semantic role labeling of speech transcripts", in Lecture Notes in Computer Science (14-20 April 2015, Cairo, Egypt), 2015, pp. 583-595.
- Simon, A., Gravier, G., Sébillot, P., Moens, M., "IRISA and KUL at MediaEval 2014: Search and hyperlinking task", in Proceedings of MediaEval benchmarking initiative for multimedia evaluation (16-17 October 2014, Barcelona, Spain), 2014, art. nr. <http://ceur-ws.org/Vol-1263/>.
- Soylu, A., "Exploiting metadata, ontologies and semantics to design/enhance new end-user experiences for adaptive pervasive computing environments.", PhD Thesis (Advisor(s): De Causmaecker, P., Duval, E., Desmet, P.), 2012. ISBN 978-90-8649-518-4.
- Soylu, A., Mödritscher, F., De Causmaecker, P., "Ubiquitous Web Navigation through Harvesting Embedded Semantic Data: A Mobile Scenario", Integrated Computer-Aided Engineering, vol. 19, no. 1, 2012, pp. 93-109.
- Speelman D., Heylen K. (2017). From dialectometry to semantics. In: Wieling M., Kroon M., Van Noord G., Bouma G. (Eds.), bookseries: Tributes, vol: 32, From Semantics to Dialectometry. Festschrift in honor of John Nerbonne, (pp. 325-334). UK: College Publications.
- Steurs, F., van der Lek-Ciudin, I., Vanallemeersch, T., "How translators work in real-life: SCATE observations", (27-28 October 2016, Brussels, Belgium), 2016.
- Stockman, C., Truyen, F., Desmet, P., "Lex-e-learn", Leuven: Acco, 2012. ISBN 9789033488863.
- Sun, M., "Constrained Non-Negative Matrix Factorization for Vocabulary Acquisition from Continuous Speech (Niet-negatieve matrix factorisatie met randvoorwaarden voor het verwerven van een woordenschat uit continue spraak)", PhD Thesis (Advisor(s): Van hamme, H.), 2012.
- Sun, M., Li, Y., Gemmeke, J., Zhang, X., "Speech enhancement under low SNR conditions via noise estimation using sparse and low-rank NMF with Kullback-Leibler divergence", IEEE Transactions on Audio Speech and Language Processing, vol. 23, no. 7, 2015, pp. 1233-1242.
- Sun, M., Van hamme, H., "Joint training of non-negative Tucker decomposition and discrete density hidden Markov models", Computer Speech and Language, vol. 27, no. 4, 2013, pp. 969-988.
- Sun, M., Van hamme, H., "Large scale graph regularized non-negative matrix factorization with l1 normalization based on Kullback-Leibler divergence", IEEE Transactions on Signal Processing, vol. 60, no. 7, 2012, pp. 3876-3880.
- Sun, M., Van hamme, H., "Tri-factorization learning of sub-word units with application to vocabulary acquisition", in Proceedings ICASSP'2012 (25-30 March 2012, Kyoto, Japan), 2012, pp. 5177-5180.
- Sun, M., Van hamme, H., Zhang, X., "WEakly supervised hmm learning for spoken word acquisition in human computer interaction with little manual effort", in Proceedings ICSP 2014 (19-23 October 2014, Hangzhou, Zhejiang, China), 2014, pp. 1341-1345.
- Sun, M., Zhang, X., Van hamme, H., "A stable approach for model order selection in nonnegative matrix factorization", Pattern Recognition Letters, vol. 54, 2015, pp. 97-102.
- Sun, M., Zhang, X., Van hamme, H., Fang Zheng, T., "Unseen noise estimation using separable deep auto encoder for speech enhancement", IEEE-acm Transactions on Audio Speech and Language Processing, vol. 24, no. 1, 2016, pp. 93-104.
- Sun, Y., Cranen, B., Gemmeke, J., Boves, L., ten Bosch, L., Doss, M., "Using sparse classification outputs as feature observations for noise-robust ASR", in Proceedings Interspeech 2012 (9-13 September 2012, Portland, Oregon, USA), 2012, pp. 1-4.
- Sun, Y., Doss, M., Gemmeke, J., Cranen, B., ten Bosch, L., Boves, L., "Combination of sparse classification and multiple layer perceptron for noise-robust ASR", in Proceedings Interspeech 2012 (9-13 September 2012, Portland, Oregon, USA), 2012, pp. 1-4.
- Sun, Y., Gemmeke, J., Cranen, B., ten Bosch, L., Boves, L., "Fusion of parametric and non-parametric approaches to noise-robust ASR", Speech Communication, vol. 56, 2014, pp. 49-62.
- Suyanto, S., Hartati, S., Harjoko, A., Van Compernelle, D., "Indonesian syllabification using a pseudo nearest neighbour rule and phonotactic knowledge", Speech Communication, vol. 85, 2016, pp. 109-118.
- Tack, A., François, T., Desmet, P., Fairon, C., "Introducing NT2Lex: A Machine-readable CEFR-graded Lexical Resource for Dutch as a Foreign Language", (10 February 2017, KU Leuven, Belgium), 2017.
- Tack, A., François, T., Ligozat, A., Fairon, C., "Evaluating Lexical Simplification and Vocabulary Knowledge for Learners of French: Possibilities of Using the FLELex Resource", in Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016) (23-28 May 2016, Portorož, Slovenia), Calzolari, N., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S. Eds., 2016, pp. 230-236.
- Tack, A., François, T., Ligozat, A., Fairon, C., "Modèles adaptatifs pour prédire automatiquement la compétence lexicale d'un apprenant de français langue étrangère", in Actes de la conférence conjointe JEP-TALN-RECITAL 2016, volume 2 : TALN (4-8 July 2016, Paris, France), 2016, pp. 221-234.

- Tack, A., François, T., Roekhaut, S., Fairon, C., "Human and Automated CEFR-based Grading of Short Answers", in Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications (7–11 September 2017, Copenhagen, Denmark), 2017, pp. 169-179.
- ter Vrugte, J., de Jong, T., Vandercruysse, S., Elen, J., Wouters, P., van Oostendorp, H., "How heterogeneous collaboration and competition interact in prevocational game-based math education", (15-17 October 2014, Leuven, Belgium), 2014.
- ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., Elen, J., "Computer game-based mathematics education: Embedded faded worked examples facilitate knowledge acquisition", Learning and Instruction, vol. 50, 2017, pp. 44-53.
- ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., Elen, J., "Faded worked examples facilitate knowledge acquisition in prevocational computer game-based mathematics education", AERA online paper repository (27 April - 1 May 2017, San Antonio (TX), USA), 2017.
- ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., Elen, J., "How competition and heterogeneous collaboration interact in prevocational game-based mathematics education", Computers and Education, vol. 89, 2015, pp. 42-52.
- ter Vrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., Verschaffel, L., Van Dooren, W., Elen, J., Vandercruysse, S., "Content integration as a factor in math-game effectiveness. In T. de Jong (chair), Game-based learning in mathematics: Research from a media comparison and value-added approach. Symposium conducted at the", EARLI 2015, Book of Abstracts (25-29 August 2015, Limassol, Cyprus), 2015, pp. 732-732.
- ter Vrugte, J., de Jong, T., Wouters, P., Vandercruysse, S., Elen, J., van Oostendorp, H., "When a game supports prevocational math education but integrated reflection does not", Journal of Computer Assisted Learning, vol. 31, no. 5, 2015, art.nr. doi 10.1111/jcal.12104, pp. 462-480.
- Thorne, S., Cornillie, F., Desmet, P., Eds, "ReCALL special issue: Digital games for language learning: challenges and opportunities", Cambridge: Cambridge University Press, 2012.
- Tjong Kim Sang, E., Bollmann, M., Boschker, R., Casacuberta, F., Dietz, F., Dipper, S., Domingo, M., van der Goot, R., van Koppen, M., Ljubevic, N., Ostling, R., Petran, F., Pettersson, E., Scherrer, Y., Schraagen, M., Sevens, L., Tiedemann, J., Vanallemeersch, T., Zervanou, K., "The CLIN27 Shared Task: Translating Historical Text to Contemporary Language for Improving Automatic Linguistic Annotation", Computational Linguistics in the Netherlands Journal, vol. 7, 2017, pp. 53-64.
- Triguero, I., Basgalupp, M., Cerri, R., Schietgat, L., Vens, C., "Partitioning the target space in multi-output learning", Proceedings of the 25th Belgian-Dutch Machine Learning Conference (Benelearn) (12-13 September 2016, Kortrijk, Belgium), 2016.
- van de Loo, J., De Pauw, G., Gemmeke, J., Karsmakers, P., Van Den Broeck, B., Daelemans, W., Van hamme, H., "Towards shallow grammar induction for an adaptive assistive vocal interface: a concept tagging approach", in Proceedings NLP4ITA 2012 (27 May 2012, Istanbul, Turkey), Rello, L., Saggion, H. Eds., 2012, pp. 27-34.
- van de Loo, J., Gemmeke, J., De Pauw, G., Driesen, J., Van hamme, H., Daelemans, W., "Towards a self-learning assistive vocal interface: vocabulary and grammar learning", in Proceedings SMIAE 2012 (12 July 2012, Jeju Island, Republic of Korea), 2012, pp. 34-42.
- Van de Velde, F., Smessaert, H., Van Eynde, F., Verbrugge, S., "Inleiding" in Patroon en argument. Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst, Van de Velde, F., Smessaert, H., Van Eynde, F., Verbrugge, S. Eds., Leuven: Universitaire Pers Leuven, 2014, ISBN 9789462700147, pp. 5-8.
- Van de Velde, F., Smessaert, H., Van Eynde, F., Verbrugge, S., Eds, "Patroon en argument. Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst", Leuven: Universitaire Pers Leuven, 2014. ISBN 9789462700147.
- Van Den Broeck, B., Bertrand, A., Karsmakers, P., Vanrumste, B., Van hamme, H., Moonen, M., "Time-domain generalized cross correlation phase transform sound source localization for small microphone arrays", in Proceedings of EDERC2012: 5th European DSP Education and Research Conference (13-14 September 2012, Amsterdam, The Netherlands), 2012.
- Van Den Broeck, B., Karsmakers, P., Van hamme, H., Vanrumste, B., "Noise robust footstep location estimation using a wireless acoustic sensor network", Journal of Ambient Intelligence and Smart Environments, vol. 8, no. 6, 2016, pp. 665-679.
- Van Den Broeck, B., Moonen, M., Van hamme, H., Vanrumste, B., "Audio based home monitoring system for the elderly", (23 March 2012, Gent, Belgium), 2012.
- Van Den Broeck, B., Vuegen, L., Dekkers, G., Van hamme, H., Moonen, M., Vanrumste, B., Karsmakers, P., "Footstep Localization By Acoustic Signals", (5-6 December 2013, Brussels, Belgium), 2013.
- Van Den Broeck, B., Vuegen, L., Van hamme, H., Moonen, M., Karsmakers, P., Vanrumste, B., "Footstep localization based on in-home microphone-array signals", in Assistive technology: From research to practice

- (AAATE2013) (19-22 September 2013, Vilamoura, Portugal), Encarnação, P., Azevedo, L., Gelderblom, G., Newell, A., Mathiassen, N. Eds., 2013, pp. 90-94.
- van der Lek-Ciudin, I., Vanallemeersch, T., De Wachter, K., "Contextual Inquiries at translators' workplace", (18-20 June 2015, Angers, France), 2015.
- van der Lek-Ciudin, I., Vanallemeersch, T., De Wachter, K., "Contextual Inquiries at translators' workplace", in TAO-CAT: proceedings (18-20 June 2015, Angers, France), Phister, B., Cancio, C., Cuzzolin, L. Eds., 2015, pp. 1-8.
- Van Eynde, F., "Predicative constructions. From the Fregean to a Montagovian treatment", "Studies in Constraint-Based Lexicalism, Stanford CA: CSLI Publications, 2015. ISBN 978-1-57586-837-0.
- Van Eynde, F., "A treebank-grounded investigation of the agreement between predicate nominals and their target", (17 January 2014, Leiden), 2014.
- Van Eynde, F., "Adposition clusters in Dutch ?!", (13 May 2016, Louvain-la-Neuve), 2016.
- Van Eynde, F., "Adposition clusters in Dutch" in Crossroad Semantics. Computation, experiment and grammar, Reckman, H., Cheng, L., Hijzelendoorn, M., Sybesma, R. Eds., Amsterdam: John Benjamins, 2017, ISBN 9789027212481, pp. 77-92.
- Van Eynde, F., "Binding, agreement and predicative complements", (10-14 August 2015, Singapore), 2015.
- Van Eynde, F., "Circumposition stranding and clustering in Dutch", (24-25 March 2017, 2017).
- Van Eynde, F., "Complement raising in Dutch", (28-30 April 2013, Stanford University), 2013.
- Van Eynde, F., "Copular constructions: from the Fregean to a Montagovian treatment", (9-10 May 2014, Seoul), 2014.
- Van Eynde, F., "How regular a pattern in these labyrinths of constructions!", (25-29 July 2016, Warsaw), 2016.
- Van Eynde, F., "NP-internal Inverted Predication", (16-17 November 2015, Frankfurt), 2015.
- Van Eynde, F., "On the agreement between predicative complements and their target", (11-12 May 2012, Frankfurt), 2012.
- Van Eynde, F., "On the agreement between predicative complements and their target", in Proceedings of the 19th International Conference on Head-driven Phrase Structure Grammar (18-21 July 2012, Chungnam), Mueller, S. Eds., 2012, pp. 348-366.
- Van Eynde, F., "Phrase Structure Grammars" in Oxford Bibliographies, New York: Oxford University Press, 2014, ISBN 9780199772810, pp. 1-37.
- Van Eynde, F., "Predicative constructions in HPSG", (23-25 June 2015, Zichron Ya'akov, Israel), 2015.
- Van Eynde, F., "Regularity and idiosyncrasy in the formation of nominals", *Journal of Linguistics*, vol. 55, 2019.
- Van Eynde, F., "Two types of predicative complements", (17-18 November 2014, 2014).
- Van Eynde, F., Augustinus, L., "Complement Raising, Extraction and Adposition Stranding in Dutch", in Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar (28-29 August 2014, Buffalo), Müller, S. Eds., 2014, pp. 156-175.
- Van Eynde, F., Augustinus, L., "Why and How to Differentiate Complement Raising from Subject Raising in Dutch", in Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar (28-29 August 2013, Berlin), Müller, S. Eds., 2013, pp. 222-242.
- Van Eynde, F., Augustinus, L., Schuurman, I., Vandeghinste, V., "Hebben of zijn bij IPP's", *Leuvense Bijdragen: Tijdschrift voor Germaanse Filologie*, vol. 99-100, 2016, pp. 11-28.
- Van Eynde, F., Augustinus, L., Schuurman, I., Vandeghinste, V., "Het verrassende resultaat van een copulativiteitspeiling" in *Patroon en Argument. Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst*, Van de Velde, F., Smessaert, H., Van Eynde, F., Verbrugge, S. Eds., Leuven: Universitaire Pers Leuven, 2014, ISBN 9789462700147, pp. 47-62.
- Van Eynde, F., Augustinus, L., Vandeghinste, V., "Number agreement in copular constructions. A treebank-based investigation", *Lingua: International Review of General Linguistics*, vol. 178, 2016, pp. 104-126.
- Van Eynde, F., Eds., "Linguistic Issues in Machine Translation", "Linguistics: Communication in Artificial Intelligence, London UK: Bloomsbury, 2015. ISBN 978-1-4742-4654-5.
- Van Eynde, F., Kim, J., "Loose Apposition: A construction-based analysis", *Functions of Language*, vol. 23, no. 1, 2016, art.nr. DOI 10.1075/fol.23.1.02kim, pp. 17-39.
- Van hamme, H., "A diagonalized Newton Algorithm for non-negative sparse coding", in Proceedings ICASSP'2013 (26-31 May 2013, Vancouver, Canada), 2013, pp. 1-5.
- Van hamme, H., "An on-line NMF model for temporal pattern learning", in *Lecture Notes in Computer Science* (12-15 March 2012, Tel Aviv, Israel), 2012, pp. 306-313.
- Van hamme, H., "The diagonalized Newton algorithm for nonnegative matrix factorization", in Proceedings international conference on learning representations (2-4 May 2013, Scottsdale, AZ, USA), 2013, pp. 1-9.

- Van Leeuwen, D., Bahari, M., "Calibration of probabilistic age recognition", in Proceedings Interspeech 2012 (9-13 September 2012, Portland, Oregon, USA), 2012, pp. 502-505.
- Van Nieuwenhuysse, K., Depaepe, F., Verschaffel, L., Wils, K., "Dealing with sources in history education. The development of a training package for prospective teachers", (29-31 August 2016, Trabzon), 2016.
- Van Nieuwenhuysse, K., Roose, H., Depaepe, F., Frère, G., Vandevenne, K., Verlinden, E., Verschaffel, L., Wils, K., "Historisch denken over bronnen. Aan de slag in het geschiedenisonderwijs" ACCO, 2016. ISBN 978-94-6292-574-8.
- Van Nieuwenhuysse, K., Roose, H., Depaepe, F., Verschaffel, L., Wils, K., "Bronnengebruik in de geschiedenisles. Redeneren met en over historische bronnen", (11 May 2016, Leuven), 2016.
- Van Nieuwenhuysse, K., Roose, H., Depaepe, F., Verschaffel, L., Wils, K., "Reasoning with and/or about sources? The use of primary sources in Flemish secondary school history education", *Historical Encounters*, vol. 4, no. 2, 2017, pp. 48-70.
- Van Nieuwenhuysse, K., Roose, H., Wils, K., Depaepe, F., Verschaffel, L., "Observation research on the use of primary sources in Flemish secondary school history classes. In C. Van Boxtel (Chair), 'Mind the gap!' Defining effective pedagogy for fostering domain-specific reasoning in history. Symposium conducted at the biennial meeting of the European Association for Research on Learning and Instruction (EARLI)", EARLI 2015, Book of Abstracts (25 - 29 August 2015, Limassol Cyprus), 2015, pp. 985-985.
- Van Nieuwenhuysse, K., Van Havere, T., Depaepe, F., Verschaffel, L., Wils, K., "Wat blijft er hangen bij leerlingen aan het einde van het secundair onderwijs? Een onderzoek naar de beeldvorming over het nationale verleden na 6 jaar geschiedenisonderwijs", (13 December 2016, Leuven), 2016.
- Van Nieuwenhuysse, K., Van Havere, T., Wils, K., Depaepe, F., Verschaffel, L., "Les jeunes, les professeurs d'histoire et le passé national. Premiers résultats d'une étude des représentations et conceptions de l'histoire nationale par des étudiants et professeurs en Belgique", (19-20 May 2014, Lyon), 2014.
- Van Nieuwenhuysse, K., Van Havere, T., Wils, K., Depaepe, F., Verschaffel, L., "Students and the national past. Research into the representation of and beliefs about the national past among students at the end of the 12th grade and compulsory history education in Flanders (Belgium)", (16-20 April 2015, Chicago), 2015.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Draye, G., Verschaffel, L., "Making the constructed nature of history visible. Flemish secondary history education through the lens of written exams" in *Joined-up history. New directions in history education research*, Chapman, A., Wilschut, A. Eds. Information Age Publishing (IAP), 2015, ISBN 978-1-68123-032-0, 978-1-68123-033-7, pp. 231-253.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Verschaffel, L., "De spanning tussen heden en verleden. Een onderzoek naar de rol van het heden bij het evalueren van het vak geschiedenis", (13 November 2013, Leuven), 2013.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Verschaffel, L., "Het heden in het Vlaamse geschiedenisonderwijs, bekeken door de bril van geschiedenisexamens", *Hermes: Tijdschrift van de Vlaamse Leraren Geschiedenis*, vol. 18, no. 55, 2014, pp. 47-69.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Verschaffel, L., "The present in Flemish secondary history education, through the lens of written history exams", *McGill Journal of Education*, vol. 50, no. 2/3, 2015, pp. 433-452.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Verschaffel, L., "The relationship between past and present in end of term tests in Flemish history education", (3-5 September 2012, Rome), 2012.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Verschaffel, L., "The relationship between past and present in end of term tests in Flemish history education", (25-28 October 2012, Québec), 2012.
- Van Nieuwenhuysse, K., Wils, K., Clarebout, G., Verschaffel, L., "The relationship between past and present in end of term tests in Flemish history education", (28-31 August 2012, Brussels), 2012.
- Van Noord, G., Bouma, G., Van Eynde, F., De Kok, D., Van der Linde, J., Schuurman, I., Tjong Kim Sang, E., Vandeghinste, V., "Large Scale Syntactic Annotation of Written Dutch: Lassy" in *Essential Speech and Language Technology for Dutch: resources, tools and applications.*, Spyns, P., Odiijk, J. Eds. Springer, 2013, ISBN 978-3-642-30909-0, 978-3-642-30910-6, pp. 147-164.
- Van Nuland, M., "Doctor-Patient Communication in the Vocational Training of General Practitioners", PhD Thesis (Advisor(s): Goedhuys, J.), (Co-Advisor(s): Van Den Noortgate, W., Heyrman, J.), 2012. ISBN 9789090266114.
- Vanallemeersch, T., "A data-driven machine translation approach using semantic tree alignment", PhD Thesis (Advisor(s): Van Eynde, F.), 2017.
- Vanallemeersch, T., "Parser-independent semantic tree alignment", (21-27 May 2012, Istanbul, Turkey), 2012.
- Vanallemeersch, T., "Parser-independent Semantic Tree Alignment", in *Proceedings of META-RESEARCH Workshop on Advanced Treebanking, in conjunction with LREC-2012* (21-27 May 2012, Istanbul, Turkey), 2012, pp. 1-5.

- Vanallemeersch, T., "Training an SRL system on cross-lingual projections in divergent parses", (20 January 2012, Tilburg, Netherlands), 2012.
- Vanallemeersch, T., Sevens, L., "Enriching machine translation input using semantics-based fuzzy matches", Book of abstracts CLIN26 (18 December 2015, Amsterdam), 2015.
- Vanallemeersch, T., Vandeghinste, V., "Assessing linguistically aware fuzzy matching in translation memories", (11-13 May 2015, Antalya, Turkey), 2015.
- Vanallemeersch, T., Vandeghinste, V., "Assessing linguistically aware fuzzy matching in translation memories", in Proceedings of EAMT 2015 (11-13 May 2015, Antalya, Turkey), 2015, pp. 153-160.
- Vanallemeersch, T., Vandeghinste, V., "Improving fuzzy matching through syntactic knowledge", (27-28 November 2014, London), 2014.
- Vanallemeersch, T., Vandeghinste, V., "Semantics-based pretranslation for SMT using fuzzy matches", in Proceedings of SSST, NAACL-HLT 2015 Workshop on Syntax, Semantics and Structure in Statistical Translation (31 May - 5 June 2015, Denver, Colorado, USA), 2015, pp. 61-64.
- Vanallemeersch, T., Vandeghinste, V., "Syntax-based fuzzy matching in translation memories", (Antwerp), 2015.
- Vandeghinste, V., "Bridging the Gap between Pictographs and Natural Language", in Proceedings of the RDWG Online Symposium on Easy-to-Read on the Web (3 december 2012, 2012).
- Vandeghinste, V., "Deyi Xiong and Min Zhang: Linguistically motivated statistical machine translation: models and algorithms", Machine Translation (previously Computers & Translation), vol. 29, no. (3), 2015, art.nr. 10.1007/s10590-015-9171-6, pp. 291-294.
- Vandeghinste, V., Augustinus, L., "Making a large treebank searchable online. The SoNaR case", in Proceedings of the LREC2014 2nd workshop on Challenges in the Management of Large Corpora (CMLC-2) (may 2014, Reykjavik), Kupietz, M., Biber, H., Lungen, H., Bański, P., Breiteneder, E., Mörth, K., Wit, A., Takhsha, J. Eds., 2014, pp. 15-20.
- Vandeghinste, V., Coppers, S., Van den Bergh, J., Vanallemeersch, T., Bulté, B., Rigouts Terryn, A., Lefever, E., van der Lek-Ciudin, I., Coninx, K., Steurs, F., "The SCATE Prototype: A Smart Computer-Aided Translation Environment", in (16-17 November 2017, London), 2017, pp. 104-113.
- Vandeghinste, V., Martens, S., Kotzé, G., Tiedemann, J., Van den Bogaert, J., De Smet, K., Van Eynde, F., Van Noord, G., "Parse and Corpus-based Machine Translation" in Essential Speech and Language Technology for Dutch: resources, tools and applications., Spyns, P., Odiijk, J. Eds. Springer, 2013, ISBN 978-3-642-30909-0, 978-3-642-30910-6, pp. 305-319.
- Vandeghinste, V., Schuurman, I., "Able-to-Include: Improving Accessibility for people with Intellectual Disabilities", in (june 2014, Dubrovnik), 2014.
- Vandeghinste, V., Schuurman, I., "Linking Pictographs to Synsets: Sclera2Cornetto", in Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) (may 2014, Reykjavik), 2014, pp. 3404-3410.
- Vandeghinste, V., Schuurman, I., Sevens, L., Van Eynde, F., "Translating text into pictographs", Natural Language Engineering, vol. 23, no. 2, 2015, pp. 217-244.
- Vandeghinste, V., Sevens, L., Schuurman, I., "E-Including the Illiterate", IEEE Potentials, vol. 36, no. 1, 2017, pp. 29-33.
- Vandeghinste, V., Van Eynde, F., "Preface", Computational Linguistics in the Netherlands Journal, vol. 7, 2017, pp. 1-2.
- Vandeghinste, V., Van Eynde, F., Eds, "Computational Linguistics in the Netherlands Journal 7", "Comoutational Linguistics in the Netherlands, vol. 7" Computational Linguistics in the Netherlands, 2017.
- Vandeghinste, V., Vanallemeersch, T., Augustinus, L., Pelemans, J., Heyman, G., van der Lek-Ciudin, I., Tezcan, A., Degraen, D., Van den Bergh, J., Macken, L., Lefever, E., Moens, M., Wambacq, P., Steurs, F., Coninx, K., Van Eynde, F., "SCATE - Smart Computer-Aided Translation Environment", Baltic Journal of Modern Computing. Special Issue. (ISSN 2255-8950) (30 May-1 June 2016, Riga, Latvia), 2016, pp. 382-382.
- Vandeghinste, V., Vanallemeersch, T., Augustinus, L., Van Eynde, F., Pelemans, J., Verwimp, L., Wambacq, P., Heyman, G., Moens, M., van der Lek-Ciudin, I., Steurs, F., Rigouts Terryn, A., Lefever, E., Tezcan, A., Macken, L., Coppers, S., Van den Bergh, J., Luyten, K., Coninx, K., "SCATE – Smart Computer-Aided Translation Environment – Year 3 (/4)", Proceedings EAMT 2017 (28-31 May 2017, Prague, Czech Republic), 2017.
- Vandeghinste, V., Vanallemeersch, T., Van Eynde, F., Heyman, G., Moens, M., Pelemans, J., Wambacq, P., van der Lek-Ciudin, I., Tezcan, A., Macken, L., Hoste, V., Geurts, E., Haesen, M., "Smart Computer Aided Translation Environment", (11-13 May 2015, Antalya, Turkey), 2015.
- Vandeghinste, V., Vanallemeersch, T., Van Eynde, F., Macken, L., Lefever, E., Hoste, V., Moens, M., Pelemans, J., Wambacq, P., Haesen, M., Coninx, K., De Wachter, K., "Smart Computer Aided Translation Environment", (16-18 June 2014, Dubrovnik, Croatia), 2014.

- Vandenbergh, B., Montero Perez, M., Reynvoet, B., Desmet, P., "Effects of form- or meaning-focused instruction on explicit and tacit lexical knowledge", (30 August - 02 September, Reading), 2017.
- Vandenbergh, B., Montero Perez, M., Reynvoet, B., Desmet, P., "Impact of form- or meaning-focused instruction on L2 vocabulary acquisition: A behavioural and neurological study", (7-8 July 2016, Nottingham), 2016.
- Vandenbergh, B., Montero Perez, M., Reynvoet, B., Desmet, P., "The Impact of Form- and Meaning-Focused Instruction on L2 Vocabulary Knowledge of Secondary School Students", (3-4 July, Reading), 2017.
- Vanderbauwhede, G., Desmet, P., "Relying data sets, tools and methods in the domain of Corpus-Based Translation Studies: three case studies French – Dutch showing the extensive possibilities of the Dutch Parallel Corpus", (29 August - 1 September 2013, Mainz University in Germersheim), 2013.
- Vanderbauwhede, G., Lauwers, P., Desmet, P., "Les emplois référentiels du SN démonstratif en français. Essai de systématisation", *Le Français Moderne*, vol. 81, no. 1, 2013, pp. 17-41.
- Vandercruysse, S., "GPS: The Game Perception Scale - Measuring students' perceptions of a GBLE", (27 – 31 August 2013, Munich, Germany), 2013.
- Vandercruysse, S., "Play your way into math! Towards an instructional design model for game-based learning in vocational secondary education", PhD Thesis (Advisor(s): Elen, J.), 2015.
- Vandercruysse, S., Clarebout, G., Elen, J., "Game-based learning: Leisure or learning? The impact of student's perceptions", (11-13 September 2012, Bari (Italy)), 2012.
- Vandercruysse, S., Clarebout, G., Elen, J., "Stimulating the development of a disposition through games-based learning", in *Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies* (4-6 July 06, Rome, Italy), 2012, pp. 730-731.
- Vandercruysse, S., Desmet, E., Vandewaetere, M., Elen, J., "Integration in the curriculum as a factor in math-game effectiveness" in *Describing and studying domain-specific serious games*, Torbeyns, J., Lehtinen, E., Elen, J. Eds., Cham - Heidelberg - New York - Dordrecht - London: Springer International Publishing AG, 2015, ISBN 978-3-319-20275-4, 978-3-319-20276-1, pp. 133-153.
- Vandercruysse, S., Desmet, E., Vandewaetere, M., Elen, J., "Integration in the curriculum as a factor in math-game effectiveness", (15-17 October 2014, Leuven, Belgium), 2014.
- Vandercruysse, S., Desmet, E., Vandewaetere, M., Verschaffel, L., Van Dooren, W., Elen, J., "Integration in the curriculum as a factor in math-game effectiveness", (27-29 August 2014, Rotterdam, the Netherlands), 2014.
- Vandercruysse, S., Elen, J., "Play your way into math! Maak het leren van verhoudingen boeiend en leuk met het wiskundespel 'Hotel Zeldenrust'", (19 May 2015, Campus Groep T, Leuven), 2015.
- Vandercruysse, S., Elen, J., "Towards a game-based learning instructional design model focusing on integration" in *Instructional techniques to facilitate learning and motivation of serious games*, Wouters, P., van Oostendorp, H. Eds., Switzerland: Springer International Publishing, 2017, ISBN 978-3-319-39296-7, pp. 17-35.
- Vandercruysse, S., Elen, J., ter Vrugte, J., Verschaffel, L., de Jong, T., Wouters, P., "The effectiveness of a math game: The impact of integrating additional support", (17-21 October 2016, Las Vegas (NV), USA), 2016.
- Vandercruysse, S., Maertens, M., Elen, J., "Description of the educational math game 'Monkey Tales: The museum of Anything'" in *Describing and studying domain-specific serious games*, Torbeyns, J., Lehtinen, E., Elen, J. Eds., Cham, Switzerland: Springer International Publishing AG., 2015, ISBN 978-3-319-20275-4, 978-3-319-20276-1, pp. 27-43.
- Vandercruysse, S., Maertens, M., Elen, J., "Description of the educational math game 'Monkey Tales: the museum of Anything'", (15-17 October 2014, Belgium: Leuven), 2014.
- Vandercruysse, S., ter Vrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., Elen, J., "Content integration as a factor in math-game effectiveness", (15-17 October 2014, Leuven, Belgium), 2014.
- Vandercruysse, S., ter Vrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., Verschaffel, L., Elen, J., "Content integration as a factor in math-game effectiveness", *Educational Technology Research and Development*, vol. 65, no. 5, 2017, pp. 1345-1368.
- Vandercruysse, S., Ter Vrugte, J., De Jong, T., Wouters, P., Van Oostendorp, H., Verschaffel, L., Elen, J., "The effectiveness of a math game: The impact of integrating additional support", (24-26 August 2016, Dijon, France), 2016.
- Vandercruysse, S., ter Vrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., Verschaffel, L., Moeyaert, M., Elen, J., "The effectiveness of a math game: The impact of integrating conceptual clarification as support", *Computers in Human Behavior*, vol. 64, 2016, pp. 21-33.
- Vandercruysse, S., ter Vrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., Verschaffel, L., Van Dooren, W., Elen, J., "'Zeldenrust': a Mathematical Game-Based Learning Environment for (pre)Vocational Students", (15-17 October 2014, Leuven, Belgium), 2014.

- Vandercruysse, S., ter Vrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., Verschaffel, L., Van Dooren, W., Elen, J., "Zeldenrust": A mathematical game-based learning environment for prevocational students" in Describing and studying domain-specific serious games, Cham - Heidelberg - New York - Dordrecht - London: Springer International Publishing AG., 2015, ISBN 978-3-319-20275-4, 978-3-319-20276-1, pp. 63-81.
- Vandercruysse, S., ter Vrugte, J., Wouters, P., Clarebout, G., Elen, J., "Development of a mathematical game-based learning environment for vocational students", (11-13 September 2012, Bari (Italy)), 2012.
- Vandercruysse, S., van Weijnen, S., Vandewaetere, M., Elen, J., "Competitie als game-element integreren in de BSO-klaspraktijk", Pedagogische Studiën, vol. 92, no. 3, 2015, pp. 179-201.
- Vandercruysse, S., Vandewaetere, M., Clarebout, G., "Game based learning: A review on the effectiveness of educational games" in Handbook of Research on Serious Games as Educational, Business, and Research Tools, Cruz-Cunha, M. Eds., Hershey, PA: IGI Global, 2012, ISBN 9781466601499, pp. 628-647.
- Vandercruysse, S., Vandewaetere, M., Cornillie, F., Clarebout, G., "Competition and students' perceptions in a game-based language learning environment", Educational Technology Research and Development, vol. 61, no. 6, 2013, pp. 927-950.
- Vandercruysse, S., Vandewaetere, M., Maertens, M., ter Vrugte, J., Wouters, P., de Jong, T., van Oostendorp, H., Elen, J., "Development and validation of the Game Perceptions Scale (GPS)", Journal of Educational Multimedia and Hypermedia, vol. 24, no. 1, 2015, pp. 43-74.
- Vandercruysse, S., Vandewaetere, M., van Weynen, S., Elen, J., "Integrating competition as a game element in the Vocational Secondary Education classroom", EARLI 2015, Book of Abstracts (25-29 August 2015, Limassol, Cyprus), 2015, pp. 102-102.
- Vandewaetere, M., Clarebout, G., "Advanced technologies for personalised learning, instruction and performance" in Handbook of research on educational communications and technology, Spector, J., Merrill, M., Elen, J., Bishop, M. Eds., New York - Heidelberg - Dordrecht - London: Springer Academic, 2014, ISBN 978-1-4614-3184-8, 978-1-4614-3185-5 (eBook), pp. 425-437.
- Vandewaetere, M., Cornillie, F., Clarebout, G., Desmet, P., "Adaptivity in Educational Games: Including Player and Gameplay Characteristics", International Journal of Higher Education, vol. 2, no. 2, 2013, pp. 106-114.
- Vandewaetere, M., Vandercruysse, S., Clarebout, G., "The Paradox of Player-Centered Adaptivity in Educational Games", (30 October - 3 November 3 2012, Louisville, Kentucky), 2012.
- Vanroy, B., Vandeghinste, V., Augustinus, L., "Querying Large Treebanks: Benchmarking GrETEL Indexing", Computational Linguistics in the Netherlands Journal, vol. 7, 2017, pp. 145-166.
- Vens, C., Saeys, Y., "Flow cytometry data mining", (10-12 March 2014, Ostend, Belgium), 2014.
- Vens, C., Verstrynge, B., Blockeel, H., "Semi-supervised clustering with example clusters", in Proceedings of the 5th International Conference on Knowledge Discovery and Information Retrieval and the 5th International Conference on Knowledge Management and Information Sharing (19-22 September 2013, Vilamoura, Portugal), Kecheng, L., Fred, A., Filipe, J. Eds., 2013, pp. 45-51.
- Verplaetse H. (2017). Differentiated impact of parallel corpus as TM for specialist and non-specialist related medical texts. Translation [and Interpreting] in Transition 3. Gent, Belgium, 13-14 July 2017.
- Verplaetse H., Heylen K. (2015). Parallel Corpora for Medical Translation Training: An Analysis of Impact On Student Performance. International Conference on Corpus Use and Learning to Translate (CULT). Alicante, Spain, 27-29 May 2015
- Verplaetse H., Lambrechts A., Heylen K. (2017). Impact of Parallel Corpora as Translation Memories on Phraseological Translation Quality in Student Translations of Specialized Medical Texts. International Conference on Corpus Linguistics, Université Paris Diderot (Sorbonne Paris Cité). Université Paris Diderot (Sorbonne Paris Cité), 31 May-2 June 2017.
- Verstraete, M., Derboven, J., Gemmeke, J., Karsmakers, P., Van Den Broeck, B., Van hamme, H., "The design of voice controlled assistive technology for people with physical disabilities", in Proceedings first international workshop on language technology in pervasive computing - LTPC 2012 (18-22 June 2012, Newcastle, UK), Anastasiou, D., Stahl, C. Eds., 2012, pp. 7-10.
- Verwimp, L., Desplanques, B., Demuynck, K., Pelemans, J., Lycke, M., Wambacq, P., "STON: Efficient subtitling in Dutch using state-of-the-art tools", in Proceedings Interspeech 2016 (8-12 September 2016, San Francisco, California, USA), 2016, pp. 780-781.
- Verwimp, L., Lahousse, K., "Definite il y a-clefts in spoken French", Journal of French Language Studies, vol. 27, no. 3, 2017, art.nr. doi:10.1017/S0959269516000132, pp. 263-290.
- Verwimp, L., Pelemans, J., Lycke, M., Van hamme, H., Wambacq, P., "Language models of spoken Dutch", arXiv:1709.03759, 2017.

- Verwimp, L., Pelemans, J., Van hamme, H., Wambacq, P., "Augmenting recurrent neural network language models with subword information", Abstract book CLIN26 (18 December 2015, Amsterdam, The Netherlands), 2015, pp. 12.
- Verwimp, L., Pelemans, J., Van hamme, H., Wambacq, P., "Character-word LSTM language models", in Proceedings EACL 2017 (3-7 April 2017, Valencia, Spain), 2017, pp. 417-427.
- Verwimp, L., Pelemans, J., Van hamme, H., Wambacq, P., "Expanding n-gram training data for language models based on morpho-syntactic transformations", Computational Linguistics in the Netherlands Journal, vol. 5, 2015, pp. 49-64.
- Verwimp, L., Pelemans, J., Van hamme, H., Wambacq, P., "Extending n-gram language models based on equivalent syntactic patterns", CLIN 2015 book of abstracts (5-6 February 2015, Antwerp, Belgium), 2015, pp. 28.
- Verwimp, L., Pelemans, J., Wambacq, P., "Project STON: speeding up Dutch subtitling with speech and language technology", Book of abstracts CLIN27 (10 February 2017, Leuven, Belgium), 2017, pp. 79.
- Virtanen, T., Gemmeke, J., Raj, B., "Active-set newton algorithm for overcomplete non-negative representations of audio", IEEE Transactions on Audio Speech and Language Processing, vol. 21, no. 11, 2013, pp. 2277-2289.
- Virtanen, T., Raj, B., Gemmeke, J., Van hamme, H., "Active-set Newton algorithm for non-negative sparse coding of audio", in Proceedings ICASSP 2014 (4-9 May 2014, Florence, Italy), 2014, pp. 3116-3120.
- Vossen, P., Maks, I., Segers, R., van der Vliet, H., Moens, M., Hofmann, K., Tjong Kim Sang, E., de Rijke, M., "Cornetto: a combinatorial lexical semantic database for Dutch" in Essential speech and language technology for Dutch: resources, tools and applications, Spyns, P., Odiijk, J. Eds., Berlin: Springer, 2012, pp. 165-184.
- Vranjes, J., Brône, G., Feyaerts, K., Paulussen, H., "Designing an online video annotation environment for language learning", (24-16 June 2013, Osnabrück, Germany), 2013.
- Vranjes, J., Brône, G., Feyaerts, K., Paulussen, H., "Implementing streaming video in academic language programs: the VIDEO.DE project", (20-23 August 2014, Groningen, The Netherlands), 2014.
- Vranjes, J., Brône, G., Feyaerts, K., Paulussen, H., "VIDEO.DE. An online video annotation tool for second language education", in 2014 CALL Conference Proceedings: Research Challenges in Call (on cd-rom) (7-9 July 2014, Antwerp, Belgium), Aerts, A. Eds., 2014, pp. 342-350.
- Vranjes, J., Brône, G., Feyaerts, K., Paulussen, H., "Video.de: Towards a student-oriented video annotation tool and language learning environment", in EDULEARN13 Proceedings (on cd-rom) (1-3 July 2013, Barcelona, Spain), Martí, D., Gómez, L., López, A. Eds., 2013, pp. 5018-5025.
- Vuegen, L., Karsmakers, P., Van hamme, H., Vanrumste, B., "ULTRASOUND BASED AUTOMATIC HOME OBSERVATION", (24-25 January 2013, Egmond aan Zee), 2013.
- Vuegen, L., Karsmakers, P., Vanrumste, B., Van hamme, H., "Acoustic Event Classification using Low-resolution Multi-label Non-negative Matrix Deconvolution", Journal of the Audio Engineering Society, no. Submitted, 2017.
- Vuegen, L., Van Den Broeck, B., Karsmakers, P., Van hamme, H., Vanrumste, B., "Automatic monitoring of activities of daily living based on real-life acoustic sensor data: A preliminary study", in Fourth workshop on speech and language processing for assistive technologies (SLPAT): Proceedings (21-22 August 2013, Grenoble, France), Alexandersson, J., Ljunglöf, P., McCoy, K., Portet, F. Eds., 2013, pp. 113-118.
- Vuegen, L., Van Den Broeck, B., Karsmakers, P., Van hamme, H., Vanrumste, B., "Energy efficient monitoring of activities of daily living using wireless acoustic sensor networks in clean and noisy conditions", in Proceedings EUSIPCO 2015 (31 August - 4 September 2015, Nice, France), 2015, pp. 449-453.
- Vuegen, L., Van Den Broeck, B., Karsmakers, P., Van hamme, H., Vanrumste, B., "Monitoring Activities of Daily Living Using Wireless Acoustic Sensor Networks in Clean and Noisy Conditions", in Proceedings EMBC 2015 (25-29 August 2015, Milano, Italy), 2015, pp. 4966-4969.
- Vulic, I., "Unsupervised Algorithms for Cross-Lingual Text Analysis, Translation Mining, and Information Retrieval (Algoritmen voor ongesuperviseerde cross-linguale tekstanalyse, het identificeren van vertalingen en informatieontsluiting)", PhD Thesis (Advisor(s): Moens, M.), (Co-Advisor(s): Crauwels, H.), 2014.
- Vulic, I., De Smet, W., Moens, M., "Cross-language information retrieval models based on latent topic models trained with document-aligned comparable corpora", Information Retrieval, vol. 16, no. 3, 2013, pp. 331-368.
- Vulic, I., De Smet, W., Tang, J., Moens, M., "Probabilistic topic modeling in multilingual settings: a short overview of its methodology with applications", in Proceedings of the NIPS Workshop on Cross-Lingual Technologies (xLiTe) (7-8 December 2012, Lake Tahoe, NV, USA), 2012, pp. 1-11.
- Vulic, I., De Smet, W., Tang, J., Moens, M., "Probabilistic topic modeling in multilingual settings: an overview of its methodology and applications", Information Processing & Management, vol. 51, no. 1, 2015, pp. 111-147.
- Vulić, I., Kiela, D., Clark, S., Moens, M., "Multi-modal representations for improved bilingual lexicon learning", in Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (7-12 August 2016, Berlin, Germany), 2016, pp. 188-194.

- Vulic, I., Moens, M., "A study on bootstrapping bilingual vector spaces from non-parallel data (and nothing else)", in Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013) (18-21 October 2013, Seattle, WA, USA), 2013, pp. 1613-1624.
- Vulic, I., Moens, M., "A unified framework for monolingual and cross-lingual relevance modeling based on probabilistic topic models", in Lecture Notes in Computer Science (24-27 March 2013, Moscow, Russian Federation), 2013, pp. 98-109.
- Vulić, I., Moens, M., "Bilingual distributed word representations from document-aligned comparable data", The Journal of Artificial Intelligence Research, vol. 55, 2016, pp. 953-994.
- Vulic, I., Moens, M., "Bilingual word embeddings from non-parallel document-aligned data applied to bilingual lexicon induction", in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (ACL 2015) (26-31 July 2015, Beijing, China), 2015, pp. 719-725.
- Vulic, I., Moens, M., "Cross-lingual semantic similarity of words as the similarity of their semantic word responses", in Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2013) (9-16 June 2013, Atlanta, GA, USA), 2013, pp. 106-116.
- Vulic, I., Moens, M., "Detecting highly confident word translations from comparable corpora without any prior knowledge", in Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012) (23-27 April 2012, Avignon, France), 2012, pp. 449-459.
- Vulic, I., Moens, M., "Monolingual and cross-lingual information retrieval models based on (bilingual) word embeddings", in Proceedings of the 38th Annual ACM SIGIR Conference on Research and Development in Information Retrieval - Full Papers (SIGIR 2015) (9-13 August 2015, Santiago, Chile), 2015, pp. 363-372.
- Vulic, I., Moens, M., "Multilingual probabilistic topic modeling and its applications in Web mining and search", in Proceedings of the 7th ACM international conference on Web search and data mining, 2014, pp. 681-682.
- Vulic, I., Moens, M., "Probabilistic models of cross-lingual semantic similarity in context based on latent cross-lingual concepts induced from comparable data", in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014) (October 25-29, 2014, Doha, Qatar), 2014, pp. 349-362.
- Vulic, I., Moens, M., "Sub-corpora sampling with an application to bilingual lexicon extraction", in Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012) (8-15 December 2012, Mumbai, India), 2012, pp. 2721-2738.
- Vulic, I., Zoghbi, S., Moens, M., "Learning to bridge colloquial and formal language applied to linking and search of e-commerce data", in Proceedings of the the 37th annual ACM SIGIR conference on research and development in information retrieval (SIGIR 2014) (6-11 July 2014, Gold Coast, Australia), 2014, pp. 1195-1198.
- VULIC, Ivan & MOENS, Marie-Francine A Unified Framework for Monolingual and Cross-Lingual Relevance Modeling Based on Probabilistic Topic Models. In Proceedings of the 35th European Conference on Information Retrieval (ECIR 2012). (Lecture Notes in Computer Science 7814) (pp. 98-109). Berlin: Springer. 2013
- Walter, O., Despotovic, V., Haeb-Umbach, R., Gemmeke, J., Ons, B., Van hamme, H., "An Evaluation of Unsupervised Acoustic Model Training for a Dysarthric Speech Interface", in Proceedings Interspeech 2014 (14-18 September 2014, Singapore), 2014, pp. 1013-1017.
- Wambacq, P., Demuynck, K., Van Compernelle, D., "SPRAAK: Speech Processing, Recognition and Automatic Annotation Kit" in Essential speech and language technology for Dutch, Spyns, P., Odiijk, J. Eds., New York, Dordrecht, London: Springer, 2013, ISBN 978-3-642-30909-0, pp. 95-113.
- Wan, H., Moens, M., Luyten, W., Zhou, X., Mei, Q., Liu, L., Tang, J., "Extracting relations from traditional Chinese medicine literature via heterogeneous entity networks", Journal of the American Medical Informatics Association, vol. 23, no. 2, 2016, art.nr. <http://dx.doi.org/10.1093/jamia/ocv092>, pp. 356-365.
- Wang, Y., Gemmeke, J., Demuynck, K., Van hamme, H., "Missing data solutions for robust speech recognition" in Essential speech and language technology for Dutch Springer, 2013, ISBN 978-3-642-30909-0, pp. 289-304.
- Wang, Y., Van hamme, H., "Multi-candidate missing data imputation for robust speech recognition", EURASIP Journal on Audio, Speech and Music Processing, vol. 17, 2012, pp. 20 pp..
- Wauters, K., "Adaptive item sequencing in item-based learning environments", PhD Thesis (Advisor(s): Van Den Noortgate, W.), (Co-Advisor(s): Desmet, P.), 2012. ISBN 978-94-6197-024-4.
- Wauters, K., Desmet, P., Van Den Noortgate, W., "Item Difficulty Estimation: an Auspicious Collaboration Between Data and Judgment", Computers and Education, vol. 58, no. 4, 2012, pp. 1183-1193.

- Wauters, K., Moeyaert, M., Desmet, P., Van Den Noortgate, W., "When easy becomes boring and difficult becomes frustrating: Disentangling the effects of item difficulty level and person proficiency on learning and motivation", *Systems*, vol. 4, no. 1, 2016, art.nr. 14.
- Weninger, F., Wöllmer, M., Geiger, J., Schuller, B., Gemmeke, J., Hurmalainen, A., Virtanen, T., Rigoll, G., "Non-negative matrix factorization for highly noise-robust ASR: to enhance or to recognize", in *Proceedings ICASSP'2012 (25-30 March 2012, Kyoto, Japan)*, 2012, pp. 4681-4684.
- Wielfaert T., Heylen K., Speelman D., Geeraerts D. (2016). Visual Analytics for Distributional Semantic Model Comparisons. In Hautli-Janisz, A. (Ed.), Lyding, V. (Ed.), *Proceedings of the LREC 2016 Workshop: VisLR II: Visualization as Added Value in the Development, Use and Evaluation of Language Resources. VISLR. Portorož, Slovenia, 23 May 2016 (art.nr. 9) (pp. 24-29) ELRA.*
- Windhouwer, M., Schuurman, I., "Linguistic resources and cats: how to use ISOcat, RELcat and SCHEMAcat", in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) (26-31 May 2014, Reykjavik)*, 2014, pp. 3806-3810.
- Windhouwer, M., Schuurman, I., Wright, S., "Collaboratively Defining Widely Accepted Linguistic Data Categories in the ISOcat Data Category Registry", in *Proceedings of ESWC2013 Workshop on Semantic Web Collaborative Spaces (27 May 2013, Montpellier)*, 2013, pp. 1-6.
- Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruyssen, S., de Jong, T., Elen, J., "The role of curiosity-triggering events in game-based learning for mathematics", (15-17 October 2014, Leuven, Belgium), 2014.
- Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruyssen, S., de Jong, T., Elen, J., "The Design and Development of the Math Game Zeldenrust", in *Proceedings of the 7th European Conference of Game Based Learning (3-4 October 2013, Porto, Portugal)*, Escudeiro, P., Vaz de Carvalho, C. Eds., 2013, pp. 191-197.
- Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruyssen, S., de Jong, T., Elen, J., "The effect of surprising events in a serious game on learning mathematics", *British Journal of Educational Technology*, vol. 48, no. 3, 2017, pp. 860-877.
- Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruyssen, S., de Jong, T., Elen, J., "The role of curiosity-triggering events in game-based learning for mathematics" in *Describing and Studying Domain-Specific Serious Games*, Torbeyns, J., Lehtinen, E., Elen, J. Eds., Cham, Switzerland: Springer International Publishing AG, 2015, ISBN 978-3-319-20275-4, 978-3-319-20276-1, pp. 191-207.
- Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruyssen, S., de Jong, T., Elen, J., "The role of surprise in game-based learning for mathematics", in *Games and Learning Alliance. 4th International Conference, GALA 2015, Rome, Italy, December 9-11, 2015, Revised Selected Papers (December 9-11, 2015, Rome, Italy)*, De Gloria, A., Veltkamp, R. Eds., 2016, pp. 401-410.
- Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruyssen, S., de Jong, T., Elen, J., "The role of surprising events in a math-game on proportional reasoning", *EARLI 2015, Book of Abstracts (25-29 August 2015, Limassol, Cyprus)*, 2015, pp. 727-727.
- Wright, S., Windhouwer, M., Schuurman, I., Broeder, D., "Segueing from a Data Category Registry to a Data Concept Registry", in *Proceedings of the 11th international conference on Terminology and Knowledge Engineering 2014 (19-21 June 2014, Berlin)*, 2014, pp. 177-187.
- Wright, S., Windhouwer, M., Schuurman, I., Kemps-Snijders, M., "Community Efforts around the ISOcat Data Category Registry" in *The People's Web Meets NLP. Collaboratively Constructed Language Resources.*, Gurevych, I., Kim, J. Eds., Berlin Heidelberg: Springer Verlag, 2013, ISBN 978-3-642-35084-9, 978-3-642-35085-6, pp. 349-373.
- Xia, H., Li, J., Tang, J., Moens, M., "Plink-LDA: Using link as prior information in topic modeling", in *Lecture Notes in Computer Science (15-19 April, 2012, Busan, South Korea)*, 2012, pp. 213-227.
- Yang, Y., Luyten, W., Liu, L., Moens, M., Tang, J., Li, J., "Forecasting potential diabetes complications", in *Proceedings of the twenty-eighth AAAI conference on artificial intelligence (AAAI-14) (27-31 July 2014, Québec city, Canada)*, 2014, pp. 313-319.
- Yilmaz, E., "Noise Robust Exemplar Matching for Speech Recognition and Enhancement", PhD Thesis (Advisor(s): Van hamme, H.), (Co-Advisor(s): Van Compernelle, D.), 2015.
- Yilmaz, E., Baby, D., Van hamme, H., "Noise robust exemplar matching for speech enhancement: Applications to automatic speech recognition", in *Proceedings Interspeech 2015 (6-10 September 2015, Dresden, Germany)*, 2015, pp. 688-692.
- Yilmaz, E., Baby, D., Van hamme, H., "Noise robust exemplar matching with coupled dictionaries for single-channel speech enhancement", in *Proceedings EUSIPCO 2015 (31 August - 4 September 2015, Nice, France)*, 2015, pp. 879-883.
- Yilmaz, E., Gemmeke, J., Van Compernelle, D., Van hamme, H., "Noise-robust digit recognition with exemplar-based sparse representations of variable length", in (23-26 September 2012, Santander, Spain), 2012.

- Yilmaz, E., Gemmeke, J., Van hamme, H., "Data selection for noise robust exemplar matching", in Proceedings ICASSP 2016 (20-25 March 2016, Shanghai, China), 2016, pp. 5980-5984.
- Yilmaz, E., Gemmeke, J., Van hamme, H., "Embedding time warping in exemplar-based sparse representations of speech", in Proceedings ICASSP'2013 (26-31 May 2013, Vancouver, Canada), 2013, pp. 8076-8080.
- Yilmaz, E., Gemmeke, J., Van hamme, H., "Exemplar selection techniques for sparse representations of speech using multiple dictionaries", in Proceedings EUSIPCO 2013 (9-13 September 2013, Marrakesh, Morocco), 2013, pp. 1-5.
- Yilmaz, E., Gemmeke, J., Van hamme, H., "Noise robust exemplar matching using sparse representations of speech", IEEE Transactions on Audio Speech and Language Processing, vol. 22, no. 8, 2014, pp. 1306-1319.
- Yilmaz, E., Gemmeke, J., Van hamme, H., "Noise robust exemplar matching with alpha-beta divergence", Speech Communication, vol. 76, 2016, pp. 127-142.
- Yilmaz, E., Gemmeke, J., Van hamme, H., "Noise-robust automatic speech recognition with exemplar-based sparse representations using multiple length adaptive dictionaries", in Proceedings CHIME 2013 (1 June 2013, Vancouver, Canada), 2013, pp. 39-43.
- Yilmaz, E., Gemmeke, J., Van hamme, H., "Noise-robust speech recognition with exemplar-based sparse representations using alpha-beta divergence", in Proceedings ICASSP 2014 (4-9 May 2014, Florence, Italy), 2014, pp. 5539-5543.
- Yilmaz, E., Pelemans, J., Lievens, S., Van hamme, H., "Speech reception threshold measurement using automatic speech recognition", in Proceedings ICASSP 2014 (4-9 May 2014, Florence, Italy), 2014, pp. 1-4.
- Yilmaz, E., Pelemans, J., Van hamme, H., "Automatic assessment of children's reading with the FLVoR decoding using a phone confusion model", in Proceedings Interspeech 2014 (14-18 September 2014, Singapore), 2014, pp. 969-972.
- Yilmaz, E., Rematas, K., Tuytelaars, T., Van hamme, H., "Learning like a toddler: watching television series to learn vocabulary from images and audio", in Proceedings ACM MM'14 (3-7 November 2014, Orlando, Florida, USA), 2014, pp. 1189-1192.
- Yilmaz, E., Van Compernelle, D., Van hamme, H., "Combining exemplar-based matching and exemplar-based sparse representations of speech", in (14 September 2012, Portland, Oregon, USA), 2012.
- Yilmaz, E., Van Compernelle, D., Van hamme, H., "Robust tracking for automatic reading tutors", in Proceedings Interspeech 2012 (9-13 September 2012, Portland, Oregon, USA), 2012, pp. 1-4.
- Yilmaz, E., Van hamme, H., Gemmeke, J., "Adaptive noise dictionary design for noise robust exemplar matching of speech", in Proceedings EUSIPCO 2015 (31 August - 4 September 2015, Nice, France), 2015, pp. 1711-1715.
- Zaman, B., Poels, Y., Sulmon, N., Annema, J., Verstraete, M., Cornillie, F., De Grooff, D., Desmet, P., "Concepts and Mechanics for Educational Mini-Games A Human-Centred Conceptual Design Approach involving Adolescent Learners and Domain Experts", International Journal On Advances in Intelligent Systems, vol. 5, no. 3-4, 2012, pp. 567-576.
- Zamith Santos, B., Gomes Mantovani, R., Schietgat, L., Vens, C., Cerri, R., "Decision trees for hierarchical classification of transposable elements", Proceedings of the 25th Belgian-Dutch Machine Learning Conference (Benelearn) (12-13 September 2016, Kortrijk, Belgium), 2016, pp. 1-3.
- Zegers, J., Van hamme, H., "Improving source separation via multi-speaker representations", in Proceedings Interspeech 2017 (20-24 August 2017, Stockholm, Sweden), 2017, pp. 1919-1923.
- Zegers, J., Van hamme, H., "Joint sound source separation and speaker recognition", in Proceedings Interspeech 2016 (8-12 September 2016, San Francisco, California, USA), 2016, art. nr. 773, pp. 2228-2232.
- Zhang, J., Tang, J., Ma, C., Tong, H., Jing, Y., Li, J., Luyten, W., Moens, M., "Fast and flexible top-k similarity search on large networks", ACM Transactions on Information Systems, vol. 36, no. 2, 2017, pp. 13:1-13:30.
- Zhang, X., "Rapid Speaker and Environment Adaptation in Automatic Speech Recognition - Part I: Parametric Normalization; Part II: Latent Variable Approaches (Snelle adaptatie aan spreker en omgeving in automatische spraakherkenning - Deel I: parametrische normalisatie; Deel II: methoden met latente variabelen)", PhD Thesis (Advisor(s): Van hamme, H.), (Co-Advisor(s): Demuyne, K.), 2014.
- Zhang, X., Demuyne, K., Van Compernelle, D., Van hamme, H., "Subspace-GMM acoustic models for under-resourced languages: feasibility study", in (7-9 May 2012, Cape Town, South Africa), 2012.
- Zhang, X., Demuyne, K., Van hamme, H., "Latent variable speaker adaptation of gaussian mixture weights and means", in Proceedings ICASSP'2012 (25-30 March 2012, Kyoto, Japan), 2012, pp. 4349-4352.
- Zhang, X., Demuyne, K., Van hamme, H., "Rapid speaker adaptation in latent speaker space with non-negative matrix factorization", Speech Communication, vol. 55, no. 9, 2013, pp. 893-908.
- Zoghbi, S., "Latent Variable Models for Language and Image Understanding in Social Media and E-Commerce Data", PhD Thesis (Advisor(s): Moens, M.), 2016.
- Zoghbi, S., Heyman, G., Gomez, J., Moens, M., "Cross-modal attribute recognition in fashion", (11 Dec 2015, 2015).

- Zoghbi, S., Heyman, G., Gomez, J., Moens, M., "Cross-modal fashion search", *Lecture Notes in Computer Science*, vol. 9517, 2016, pp. 367-373.
- Zoghbi, S., Heyman, G., Gomez, J., Moens, M., "Fashion meets computer vision and NLP at e-commerce search", *International Journal of Computer and Electrical Engineering*, vol. 8, no. 1, 2016, pp. 31-43.
- Zoghbi, S., Vulic, I., Moens, M., "Are words enough? A study on text-based representations and retrieval models for linking pins to online shops", in *Proceedings of the CIKM 2013 international workshop on mining unstructured big data using natural language processing (MNLP 2013)* (October 27 - November 1, 2013, San Francisco, CA, USA), 2013, pp. 45-52.
- Zoghbi, S., Vulic, I., Moens, M., "I pinned it. Where can I buy one like it? Automatically linking Pinterest pins to online webshops", in *Proceedings of the CIKM 2013 international workshop on data-driven user behavioral modelling and mining from social media (DUBMOD 2013)* (October 27 - November 1, 2013, San Francisco, CA, USA), 2013, pp. 9-12.
- Zoghbi, S., Vulic, I., Moens, M., "Latent Dirichlet allocation for linking user-generated content and e-commerce data", *Information Sciences*, vol. 367, no. C, 2016, pp. 573-599.

B. UNIVERSITEIT ANTWERPEN - OVERZICHT PUBLICATIES

- Amado Laurel Maria Herminia, Cabral Maria de Jesus, Schuerewegen Franc, "Lire de près, de loin: close vs distant reading", *Classiques Garnier*, Paris, *Théorie littéraire*, 80, 2014, 202p.
- Beckett Samuel, Van Hulle Dirk, Weller Shane, Neyt Vincent, "The making of Samuel Beckett's *L'innommable* / *The unnamable*", Brussels, ASP, 2013
- Beloborodova Olga, "The modernist mind extended: Samuel Beckett, his fiction and the extended mind theory", *Manuscrita: revista de crítica genética*, vol. 28, 2015, p. 74-87
- Bernaerts Lars, Van Hulle Dirk, "Narrative across versions: narratology meets genetic criticism", *Tel Aviv, Poetics today*, vol. 34, nr. 3, 2013, p. 281-326
- Birnbaum David J., Bonde Sheila, Kestemont Mike, "The digital Middle Ages, Cambridge, *Speculum: a journal of medieval studies*", 92,S1 (Themanummer), 2017, 360 p.
- Birnbaum David J., Bonde Sheila, Kestemont Mike, "The digital Middle Ages: a *speculum* supplement", Cambridge, Mass., *Speculum: a journal of medieval studies*, vol. 92, nr. 4, 2017, p. 961-962
- Birnbaum David J., Bonde Sheila, Kestemont Mike, "The digital Middle Ages: a *speculum* supplement", Chicago, Univ chicago press, *Speculum: a journal of medieval studies*, vol. 92, nr. 4, 2017, p. 961-962
- Birnbaum David J., Bonde Sheila, Kestemont Mike, "The digital Middle Ages: an introduction", Chicago, Univ chicago press, *Speculum: a journal of medieval studies*, vol. 92: s:[1], 2017, S1-S38
- Canales Lea, Daelemans Walter, Boldrini Ester, Martinez-Barco Patricio, "Towards the improvement of automatic emotion pre-annotation with polarity and subjective information", *Proceedings of Recent Advances in Natural Language Processing*, Varna, Bulgaria, September 4-6 2017, 2017, p. 157-163
- Cassani Giovanni, Grimm Robert, Daelemans Walter, Gillis Steven, "Which distributional cues help the most? Unsupervised context selection for lexical category acquisition", Lisbon, Association for Computational Linguistics, *Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning / Berwick, Robert [edit.], et al.*, 2015, p. 33-39
- Cassani Giovanni, Grimm Robert, Gillis Steven, Daelemans Walter, "Constraining the search space in cross-situational word learning: different models make different predictions", Austin, Tex., Cognitive Science Society, *CogSci 2016: Recognizing and Representing Events: Integrating Psychological, Philosophical, Linguistic, Computational and Neural Perspectives: proceedings of the 38th Annual Conference of the Cognitive Science Society*, August 10-13, 2016, Philadelphia, P, 2016, p. 1152-1157
- Cassani Giovanni, Grimm Robert, Gillis Steven, Daelemans Walter, "Distributional learning and lexical category acquisition: what makes words easy to categorize?", Austin, Tex., Cognitive Science Society, *Proceedings of the 39th Annual Conference of the Cognitive Science Society / Gunzelmann, G. [edit.], et al.*, 2017, p. 216-221

- Crowley Ronan, Van Hulle Dirk, "New quotatoes: Joycean exegesis in the digital age", Leiden, Brill, *European Joyce studies*, 25, 2016, 245 p.
- Daelemans Walter, De Clercq Orphée, Hoste Véronique, "Stylene: an environment for stylometry and readability research for Dutch", London, Ubiquity Press, *CLARIN in the Low Countries / Odijk, J. [edit.], et al., 2017*, p. 195-209
- Daelemans Walter, "A tourist in the AI-lab before the revolution", Schoten, Berger & Montag Publishers, *Luc Steels: growing seeds / Beuls, Katrien [edit.], et al., 2017*, p. 20-23
- Daelemans Walter, "Explanation in computational stylometry", Berlin, Springer, *Lecture notes in computer science*, 2013, p. 451-462
- Daelemans Walter, Hoste Véronique, "STYLENE: an environment for stylometry and readability research for Dutch", Antwerpen, CLIPS, *CLiPS technical report series, CTRS-004*, 2013, 28 p.
- Daelemans Walter, Michael P. Oakes, "Literary Detective Work on the Computer", Oxford, Oxford univ press, *Digital scholarship in the humanities: a journal of the Alliance of Digital Humanities Organizations*, vol. 32, nr. 1, 2017, p. 222-223
- Daelemans Walter, van Asch Vincent, "Predicting the effectiveness of self-training: application to sentiment classification", Antwerpen, 2016, 9 p.
- Dahlström Mats, Dillen Wout, "Litteraturbanken: the Swedish literature bank", *Digital text collections / Henny-Krahmer, Ulrike [edit.]*, vol. 6, 2017, p. 1-14
- Damen David, Luyckx Kim, Hellebaut Geert, van den Bulcke Tim, "PASTEL: a semantic platform for assisted clinical trial patient recruitment", S.I., IEEE, 2013 IEEE International Conference on Healthcare Informatics (ICHI), 2013, p. 269-276
- de Bruijn Elisabeth, Kestemont Mike, "Contrastive multivariate analyses of the Middle Low German Flos unde Blankeflos tradition", Helsinki, *Neuphilologische Mitteilungen*, vol. 114, nr. 2, 2013, p. 171-205
- De Pauw Guy, de Schryver G.-M., Forcada M.L., Sarasola K., Tyers F.M., "Proceedings of the Workshop on Language Technology for Normalisation of Less-Resourced Languages (SALTMIL8/AfLaT2012)", Istanbul, European Language Resources Association (ELRA), 2012
- De Pauw Guy, de Schryver G.-M., van de Loo Janneke, "Resource-light Bantu part-of-speech tagging", Istanbul, European Language Resources Association (ELRA), *Proceedings of the Workshop on Language Technology for Normalisation of Less-Resourced Languages (SALTMIL8/AfLaT2012) / Pauw, De, G. [edit.], e.a., 2012*, p. 85-92
- De Pauw Guy, Verhoeven Ben, Desmet Bart, Lefever Els, "Proceedings of the LREC 2016 Workshop on Text Analytics for Cybersecurity and Online Safety (TA-COS 2016)", Portorož, European Language Resources Association (ELRA), 2016, 40 p.
- De Smedt Tom, de Bleser Frederik, van Asch Vincent, Nijs Lucas, Daelemans Walter, "Gravital: natural language processing for computer graphics", Berlin, De Gruyter, *Applications of cognitive linguistics*, 2013, p. 81-97
- De Smedt Tom, "Modeling creativity: case studies in Python", Antwerpen, Universiteit Antwerpen, *Faculteit Letteren en Wijsbegeerte, Departement Taalkunde*, 2013, 165 p.
- De Smedt Tom, Daelemans Walter, "Pattern for Python", Cambridge, Mass., *Journal of machine learning research*, vol. 13, 2012, p. 2063-2067
- De Smedt Tom, Daelemans Walter, "Vreselijk mooi! (terribly beautiful): a subjectivity lexicon for Dutch adjectives", Budapest, -, *Across languages and cultures*, 2012, p. 3568-3572
- De Smedt Tom, Marfia Fabio, Matteucci Matteo, Daelemans Walter, "Using Wiktionary to build an Italian part-of-speech tagger", *Lecture notes in computer science*, 8455, 2014, p. 1-8
- De Smedt Tom, Menschaert Lieven, "VALENCE: affective visualisation using EEG", Lisse, *Digital creativity*, vol. 23, nr. 3-4, 2012, p. 272-277
- De Smedt Tom, Nijs Lucas, Daelemans Walter, "Creative web services with pattern", *Proceedings of the Fifth International Conference on Computational Creativity*, Ljubljana, Slovenia, 9-13 June 2014, 2014, p. 1-3
- Dekker Ronald Haentjens, Van Hulle Dirk, Middell Gregor, Neyt Vincent, van Zundert Joris, "Computer-supported collation of modern manuscripts: CollateX and the Beckett Digital Manuscript Project", Oxford, *Literary and linguistic computing*, vol. 30, nr. 3, 2015, p. 452-470
- Dillen Wout, "(Hiatus in MS.): towards a TEI compliant typology of textual lacunae in Samuel Beckett's manuscripts", *ManuscrÝtica*, vol. 28, 2015, p. 65-73
- Dillen Wout, "A lexicon of scholarly editing", www.bookhistorynetwork.wordpress.com, 2015
- Dillen Wout, "Bridging the Babel of textual criticism: lexicographical approaches to encompassing a multilingual research field", www.fontegaia.hypotheses.org, 2016
- Dillen Wout, "Digital scholarly editing and memory institutions", www.dixit.hypotheses.org, 2016

- Dillen Wout, "Digital scholarly editing for the genetic orientation: the making of a genetic edition of Samuel Beckett's works", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departement Letterkunde, 2015, 275 p.
- Dillen Wout, "Patrick Sahle: Digitale Editionsformen: zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels", München, Bibliothek: Forschung und Praxis, vol. 38, nr. 1, 2014, p. 156-160
- Dillen Wout, "Sequentiality in genetic digital scholarly editions: models for encoding the dynamics of the writing process", Kraków, Digital Humanities 2016: conference abstracts: 11-16 July, 2016, Kraków, Poland, 2016, p. 174-175
- Dillen Wout, Macé Caroline, Van Hulle Dirk, "Texts beyond borders: multilingualism and textual scholarship", Amsterdam, Rodopi, Variants: the journal of the European Society for Textual Scholarship, 9, 2012, 296 p.
- Dillen Wout, Neyt Vincent, "Digital scholarly editing within the boundaries of copyright restrictions", Oxford, Oxford univ press, Digital scholarship in the humanities: a journal of the Alliance of Digital Humanities Organizations, vol. 31, nr. 4, 2016, p. 785-796
- Dillen Wout, Spadini Elena, Zanardo Monica, "Il Lexicon of Scholarly Editing: una bussola nella Babele delle tradizioni filologiche", Roma, Carocci, Ecdotica, vol. 13, 2016, p. 169-19
- Eder M., Rybicki J., Kestemont Mike, "Stylometry with R: a package for computational text analysis", Boston, MA, R Foundation for Statistical Computing, The R journal, vol. 8, nr. 1, 2016, p. 107-121
- Eder Maciej, Rybicki Jan, Kestemont Mike, "Stylometry with R: a suite of tools", S.I., Center for Digital Research in the Humanities, Digital Humanities 2013: conference abstracts, University of Nebraska, Lincoln, USA, 16-19 July 2013, 2013, p. 487-489
- Emmery Chris, Chrupala Grzegorz, Daelemans Walter, "Simple queries as distant labels for detecting gender on Twitter", Copenhagen, Association for Computational Linguistics, Proceedings of the 3rd Workshop on Noisy User-generated Text, 2017, p. 50-55
- Fivez Pieter, Suster Simon, Daelemans Walter, "Unsupervised context-sensitive spelling correction of clinical free-text with word and character n-gram embedding", Association for Computational Linguistics, 16th Workshop on Biomedical Natural Language Processing of the Association for Computational Linguistics, 2017, p. 143-148
- Fivez Pieter, Suster Simon, Daelemans Walter, "Unsupervised context-sensitive spelling correction of English and Dutch clinical free-text with word and character n-gram embeddings", Computational Linguistics in the Netherlands Journal, nr. 7, 2017, p. 39-52
- Gemmeke Jort F., van de Loo Janneke, De Pauw Guy, Driesen Joris, Van Hamme Hugo, Daelemans Walter, "A self-learning assistive vocal interface based on vocabulary learning and grammar induction", Proceedings of INTERSPEECH 2012, 2012, p. 1-4
- Gemmeke Jort, Ons Bart, Tessema Netsanet, Van hamme Hugo, van de Loo Janneke, De Pauw Guy, Daelemans Walter, Huyghe Jonathan, Derboven Jan, Vuegen Lode, Van Den Broeck Bert, Karsmakers Peter, Vanrumste Bart, "Self-taught assistive vocal interfaces: an overview of the ALADIN project", Lyon, ISCA, Proceedings of Interspeech, 2013, p. 2038-2043
- Grimm Robert, Cassani Giovanni, Daelemans Walter, Gillis Steven, "Towards a model of prediction-based syntactic category acquisition: first steps with word embeddings", Lisbon, Association for Computational Linguistics, Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning / Berwick, Robert [edit.], et al., 2015, p. 28-32
- Grimm Robert, Cassani Giovanni, Gillis Steven, Daelemans Walter, "Evidence for a facilitatory effect of multi-word units on child word learning", Austin, Tex., Cognitive Science Society, Proceedings of the 39th Annual Conference of the Cognitive Science Society / Gunzelmann, G. [edit.], et al., 2017, p. 433-438
- Grimm Robert, Cassani Giovanni, Gillis Steven, Daelemans Walter, "Facilitatory effects of multi-word units in lexical processing and word learning: a computational investigation", Pully, Switzerland, Frontiers Research Foundation, Frontiers in psychology, vol. 8, 2017, 17 p.
- Guerrero Iván, Verhoeven Ben, Barbieri Francesco, Martins Pedro, Pérez y Pérez Rafael, "TheRiddlerBot: a next step on the ladder towards computational creativity", Park City, Utah, Brigham Young University, Proceedings of the Sixth International Conference on Computational Creativity / Toivonen, Hannu [edit.], et al., 2015, p. 315-322
- Haverals Wouter, de Bruijn Elisabeth, Caers Bram, Fraeters Veerle, Kestemont Mike, Stoop Patricia, Verschueren Anke, "Met een MOOC'je in een hoekje: Middelnederlandse literatuur voor een ruim publiek", 's-Gravenhage, Neerlandia: algemeen-Nederlands tijdschrift, vol. 121, 2017, p. 4-8
- Hendrickx Iris, Bouma Gosse, Daelemans Walter, Hoste Véronique, "COREA: coreference resolution for extracting answers for Dutch", Berlin, Springer, Essential speech and language technology for Dutch: results by the STEVIN-programme / Spyns, Peter [edit.], e.a., 2013, p. 97-108

- Hilte Lisa, Vandekerckhove Reinhild, Daelemans Walter, "Expressiveness in Flemish online teenage talk: a corpus-based analysis of social and medium-related linguistic variation", Ljubljana, Academic Publishing Division of the Faculty of Arts of the University of Ljubljana, Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities / Fišer, Darja [edit.], et al., 2016, p. 30-33
- Hilte Lisa, Vandekerckhove Reinhild, Daelemans Walter, "Modeling non-standard language use in adolescents' CMC: the impact and interaction of age, gender and education", Bolzano, Eurac Research, Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora17), 3-4 October 2017, Italy / Stemle, Egon W. [edit.], et al., 2017, p. 11-15
- Hogenbirk Marjolein, Kestemont Mike, "Het werk van een klerk: een stylometrische verkenning van het auteurschap achter 'Der ystorien bloeme'", Leiden, Tijdschrift voor Nederlandse taal- en letterkunde, vol. 129, nr. 4, 2013, p. 246-273
- Hürlimann Manuela, Benno Weck, van den Berg Esther, Suster Simon, Nissim Malvina, "GLAD: Groningen Lightweight Authorship Detection: Notebook for PAN at CLEF 2015", Proceedings of CLEF 2015 Labs and Workshops, Notebook Papers, CEUR Workshop, 2015, p. 1-12
- Junqué de Fortuny Enric, De Smedt Tom, Martens David, Daelemans Walter, "Design and evaluation of empirical models for stock price prediction", Antwerp, Research paper / UA, Faculty of Applied Economics, 2012:17, 2012, 22 p
- Junqué de Fortuny Enric, De Smedt Tom, Martens David, Daelemans Walter, "Evaluating and understanding text-based stock price prediction models", Oxford, Information processing and management, vol. 50, nr. 2, 2014, p. 426-441
- Junqué de Fortuny Enric, De Smedt Tom, Martens David, Daelemans Walter, "Media coverage in times of political crisis: a text mining approach", New York, Expert systems with applications, vol. 39, nr. 14, 2012, p. 11616-11622
- Junqué de Fortuny Enric, Martens David, De Smedt Tom, Daelemans Walter, "Media coverage in times of political crisis: a text mining approach", Antwerp, UA, Research paper / UA, Faculty of Applied Economics, 2012:003, 2012, 27 p.
- Kapociūtė-Dzikiėnė J., Vaassen Frederik, Daelemans Walter, Krupavicius A., "Improving topic classification for highly inflective languages", Mumbai, Proceedings of COLING 2012, the 24th International Conference on Computational Linguistics, 2012, p. 1393-1410
- Karsdorp F., Kestemont Mike, Schöch C., Van den Bosch A., "The Love Equation: computational modeling of romantic relationships in French classical drama", Atlanta, Ga, Proceedings of the 6th Workshop on Computational Models of Narrative (CMN-15), 2015, p. 98-107
- Karsdorp Folgert, Manjavacas Arévalo Enrique, Burtenshaw Benjamin, Kestemont Mike, "Synthetische literatuur: neurale netwerken als coauteur", Dixit: tijdschrift over taal- en spraaktechnologie, vol. 14, 2017, p. 18-19
- Kemps-Snijders Marc, Schuurman Ineke, Daelemans Walter, Demuyne Kris, Desplanques Brecht, Hoste Véronique, huijbregts Marijn, Martens Jean-Pierre, Paulussen Hans, Pelemans Joris, Reynaert Martin, Vandeghinste Vincent, van den Bosch Antal, et al., "TTNWW to the Rescue: No Need to Know How to Handle Tools and Resources", London, Ubiquity Press, CLARIN in the Low Countries / Odijk, J. [edit.], et al., 2017, p. 83-93
- Kestemont Mike, de Gussem J., "Integrated sequence tagging for medieval Latin using deep representation learning", Journal of Data Mining and Digital Humanities, 2017, p. 1-17
- Kestemont Mike, "A computational analysis of the scribal profiles in two of the oldest manuscripts of Hadewijch", Bruxelles, Scriptorium: revue internationale des études relatives aux manuscrits, vol. 69, 2015, p. 159-177
- Kestemont Mike, "A quantitative study of the rhyme words in the middle Dutch Arthurian epic", Berlin, Zeitschrift für deutsches Altertum und deutsche Literatur, vol. 142, nr. 1, 2013, p. 1-33
- Kestemont Mike, "A shallow reading of the Limburg sermons", Antwerpen, Ruusbroecgenootschap, Ons geestelijk erf: driemaandelijks tijdschrift voor de geschiedenis van de vroomheid in de Nederlanden, vol. 85, nr. 3, 2014, p. 145-162
- Kestemont Mike, "Auteursherkenning met rijmwoorden in de middel nederlandse Artur- en Karelepiek: eerherstel voor Icarus?", Leiden, Tijdschrift voor Nederlandse taal- en letterkunde, vol. 128, nr. 2, 2012, p. 135-159
- Kestemont Mike, "Beta humanities: de opmars van digitale technieken in de geesteswetenschappen", 's-Gravenhage, Neerlandia: algemeen-Nederlands tijdschrift, vol. 120, nr. 3, 2016, p. 4-7
- Kestemont Mike, "Computers leren taal uit het niets", Kortrijk-Heule, UGA, Over taal: tijdschrift over taal, tekst en communicatie, vol. 55, nr. 1, 2016, p. 9-11
- Kestemont Mike, "EACL toont ook de muziek van de literatuur", E-Data & Research, vol. 8, nr. 3, 2014, p. 2-2

- Kestemont Mike, "Een 'assonantic revival'? Een kwantitatief diachroon onderzoek naar de assonantie in de Middelnederlandse ridderepiek, met bijzondere aandacht voor de veertiende eeuw", *Hilversum, Verloren, Middeleeuwse studies en bronnen*, 131, 2012, p. 259-290
- Kestemont Mike, "Function words in authorship attribution: from black magic to theory?", Göteborg, Association for Computational Linguistics, Third Computational Linguistics for Literature Workshop, co-located with EACL 2014 the 14th Conference of the European Chapter of the Association for Computational Linguistics (27 April 2014, Gothenburg, Sweden), 2014, p. 59-66
- Kestemont Mike, "Het gewicht van de auteur: een onderzoek naar stylometrische auteursherkenning in de Middelnederlandse epiek", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departementen Taal- en Letterkunde, 2012, 323 p.
- Kestemont Mike, "Het gewicht van de auteur: stylometrische auteursherkenning in Middelnederlandse literatuur", Gent, Koninklijke Academie voor Nederlandse Taal- en Letterkunde, *Studies op het gebied van de oudere Nederlandse letterkunde*, 5, 2013, 324 p.
- Kestemont Mike, "In het land van Alena: over de lokalisering van de Ferguut", Münster, Center for Digital Research in the Humanities, *Want hi verkende dien name wale: opstellen voor Willem Kuiper / Hogenbirk, M. [edit.], e.a.*, 2014, p. 93-98
- Kestemont Mike, "Keizerlijk Vlaanderen: streekliteratuur in Rijks-Vlaanderen en de casus van de Karelepiek", Hilversum, *Verloren, Queeste: tijdschrift over middeleeuwse letterkunde in de Nederlanden*, vol. 19, nr. 2, 2012, p. 91-111
- Kestemont Mike, "Matthew L. Jockers: Macroanalysis: digital methods and literary history", Oxford, *Literary and linguistic computing*, vol. 29, nr. 2, 2014, p. 274-276
- Kestemont Mike, "Produits de terroir? La littérature régionale dans la Flandre impériale et le cas de la chanson de geste moyen-néerlandaise", Bâle, *Publication du Centre européen d'études bourguignonnes (14e-16e s.)*, vol. 54, 2014, p. 37-55
- Kestemont Mike, "Serendipiteit en macro-medioneerlandistiek", Utrecht, *Madoc: tijdschrift voor medievistiek*, vol. 28, nr. 3, 2014, p. 157-165
- Kestemont Mike, "Stylometric authorship attribution for the Middle Dutch mystical tradition from Groenendaal", London, *Dutch crossing: a journal of low countries studies*, 2016, p. 1-35
- Kestemont Mike, "Stylometry for medieval authorship studies: an application to rhyme words", *Digital philology*, vol. 1, nr. 1, 2012, p. 42-72
- Kestemont Mike, "What can stylometry learn from its application to Middle Dutch literature?", Amsterdam, Amsterdam University Press, *Journal of Dutch literature*, vol. 2, nr. 2, 2012, p. 46-65
- Kestemont Mike, Christlein Vincent, Stutzmann Dominique, "Artificial paleography: computational approaches to identifying script types in medieval manuscripts", Cambridge, Mass., *Speculum: a journal of medieval studies*, vol. 92: s:[1], 2017, S86-S109
- Kestemont Mike, Daelemans Walter, Sandra Dominiek, "Robust rhymes? The stability of authorial style in medieval narratives", Amsterdam, *Journal of quantitative linguistics*, vol. 19, nr. 1, 2012, p. 54-76
- Kestemont Mike, Daelemans Walter, Sandra Dominiek, "Robust rhymes? The stability of authorial style in medieval narratives", Amsterdam, *Journal of quantitative linguistics*, vol. 19, nr. 1, 2012, p. 54-76
- Kestemont Mike, De Pauw Guy, van Nie Renske, Daelemans Walter, "Lemmatization for variation-rich languages using deep learning You have access", Oxford, Oxford University Press, *Digital scholarship in the humanities: a journal of the Alliance of Digital Humanities Organizations*, vol. 32, nr. 4, 2017, p. 797-815
- Kestemont Mike, De Pauw Guy, van Nie Renske, Daelemans Walter, "Lemmatization for variation-rich languages using deep learning You have access", Oxford, Oxford University Press, *Digital scholarship in the humanities: a journal of the Alliance of Digital Humanities Organizations*, vol. 32, nr. 4, 2017, 797-815
- Kestemont Mike, Karsdorp Folgert, Düring Marten, "Mining the twentieth century's history from the time magazine corpus", Göteborg, Association for Computational Linguistics, *Proceedings of LaTeCH 2014, The 8th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, co-located with EACL 2014, the 14th Conference of the European Chapter of the Association for Computational Linguistics (26 April 2, 2014, p. 62-70
- Kestemont Mike, Luyckx Kim, Daelemans Walter, Crombez Thomas, "Cross-genre authorship verification using unmasking", Amsterdam, *English studies: a journal of English language and literature*, vol. 93, nr. 3, 2012, p. 340-356
- Kestemont Mike, Luyckx Kim, Daelemans Walter, Crombez Thomas, "Cross-genre authorship verification using unmasking", Amsterdam, *English studies: a journal of English language and literature*, vol. 93, nr. 3, 2012, p. 340-356

- Kestemont Mike, Luyckx Kim, Daelemans Walter, Crombez Thomas, "Evaluating unmasking for cross-genre authorship verification", Hamburg, Digital humanities / Meister, Jan Christoph [edit.], 2012, p. 249-252
- Kestemont Mike, Luyckx Kim, Daelemans Walter, Crombez Thomas, "Evaluating unmasking for cross-genre authorship verification", Hamburg, Digital humanities / Meister, Jan Christoph [edit.], 2012, p. 249-252
- Kestemont Mike, Moens Sara, Deploige Jeroen, "Collaborative authorship in the twelfth century: a stylometric study of Hildegard of Bingen and Guibert of Gembloux", Oxford, Literary and linguistic computing, vol. 30, nr. 2, 2015, p. 199-224
- Kestemont Mike, Moens Sara, Deploige Jeroen, "Stylometry and the complex authorship in Hildegard of Bingen's oeuvre", S.l., Center for Digital Research in the Humanities, Digital Humanities 2013: conference abstracts, University of Nebraska, Lincoln, USA, 16-19 July 2013, 2013, p. 255-258
- Kestemont Mike, Peersman Claudia, De Decker Benny, De Pauw Guy, Luyckx Kim, Morante Roser, Vaassen Frederik, van de Loo Janneke, Daelemans Walter, "The Netlog corpus: a resource for the study of Flemish Dutch internet language" Istanbul, European Language Resources Association, Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12) / Calzolari, Nicoletta [edit.], e.a., 2012, p. 1569-1572
- Kestemont Mike, Peersman Claudia, De Decker Benny, De Pauw Guy, Luyckx Kim, Morante Roser, Vaassen Frederik, van de Loo Janneke, Daelemans Walter, "The Netlog corpus: a resource for the study of Flemish Dutch internet language", Istanbul, European Language Resources Association, Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12) / Calzolari, Nicoletta [edit.], e.a., 2012, p. 1569-1572
- Kestemont Mike, Stover J., Koppel M., Karsdorp K., Daelemans Walter, "Authorship Verification with the Ruzicka Metric", Kraków, Proceedings of Digital Humanities 2016, 2016, p. 246-249
- Kestemont Mike, Stover J., Koppel M., Karsdorp K., Daelemans Walter, "Authorship Verification with the Ruzicka Metric", Kraków, Proceedings of Digital Humanities 2016, 2016, p. 246-249
- Kestemont Mike, Stover Justin, Koppel Moshe, Karsdorp Folgert, Daelemans Walter, "Authenticating the writings of Julius Caesar", New York, Expert systems with applications, vol. 63, 2016, p. 86-96
- Kestemont Mike, Stover Justin, Koppel Moshe, Karsdorp Folgert, Daelemans Walter, "Authenticating the writings of Julius Caesar", New York, Expert systems with applications, vol. 63, 2016, p. 86-96
- Kestemont Mike, Stronks Els, de Bruin Martine, de Winkel Tim, "Did a poet with donkey ears write the oldest anthem in the world? Ideological implications of the computational attribution of the Dutch national anthem to Petrus Dathenus", Digital humanities 2017. Book of abstracts (Montreal, August 9, 2017), 2017, p. 280-282
- Kestemont Mike, Stronks Els, De Bruin Martine, De Winkel Tim, "Van wie is het Wilhelmus? Auteurskwesties rond het Nederlandse volkslied met de computer onderzocht", Amsterdam, Amsterdam University Press, Meertens Nieuwjaarsuitgaven, 2017, 128 p.
- Kestemont Mike, Stutzmann Dominique, "Script identification in medieval Latin manuscripts using convolutional neural networks", Digital humanities 2017. Book of abstracts (Montreal, August 10, 2017), 2017, p. 283-285
- Laaha Sabine, Blineder Michaela, Gillis Steven, "Noun plural production in preschoolers with early cochlear implantation: an experimental study of Dutch and German", Amsterdam, International journal of pediatric otorhinolaryngology, vol. 79, nr. 4, 2015, p. 561-569
- Luyckx Kim, Vaassen Frederik, Peersman Claudia, Daelemans Walter, "Fine-grained emotion detection in suicide notes: a thresholding approach to multi-label classification", Biomedical Informatics Insights, nr. 5, 2012, p. 61-69
- Manjavacas Arévalo Enrique, de Gussem Ewoud, Daelemans Walter, Kestemont Mike, "Assessing the stylistic properties of neurally generated text in authorship attribution", Proceedings of the Workshop on Stylistic Variation, collocated with EMNLP 2017 (Copenhagen, September 8, 2017) / Association for Computational Linguistics, 2017, p. 116-125
- Manjavacas Arévalo Enrique, de Gussem Ewoud, Daelemans Walter, Kestemont Mike, "Assessing the stylistic properties of neurally generated text in authorship attribution", Proceedings of the Workshop on Stylistic Variation, collocated with EMNLP 2017 (Copenhagen, September 8, 2017) / Association for Computational Linguistics, 2017, p. 116-125
- Manjavacas Arévalo Enrique, Karsdorp Folgert, Burtenshaw Benjamin, Kestemont Mike, "Synthetic literature: writing science fiction in a co-creative process", Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017), collocated with INLG 2017 (Santiago de Compostella, September 4, 2017) / Association for Computational Linguistics, 2017, p. 29-37
- Milota Megan, "But every man cannot be a surgeon': Elizabeth Gaskell's many-sided medical practitioners", Copenhagen, Orbis litterarum, vol. 68, nr. 6, 2013, p. 473-505

- Milota Megan, "From 'compelling and mystical' to makes you want to commit suicide: quantifying the spectrum of online reader responses", Amsterdam, Scientific study of literature, vol. 4, nr. 2, 2015, p. 178-195
- Morante Roser, Daelemans Walter, "Annotating modality and negation for a machine reading evaluation", CLEF 2012 Conference and Labs of the Evaluation Forum - Question Answering For Machine Reading Evaluation (QA4MRE), Rome / Forner, P. [edit.], e.a., 2012, p. 1-11
- Morante Roser, Daelemans Walter, "ConanDoyle-neg: annotation of negation in Conan Doyle stories", European Language Association, Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC12), Istanbul, 2012, p. 1563-1568
- Morante Roser, Krallinger M., Valencia , Daelemans Walter, "Machine reading of biomedical texts about Alzheimer's disease", CLEF 2013 Evaluation Labs and Workshop: online workingnotes / Forner, P. [edit.], e.a., 2013
- Morante Roser, Krallinger Martin, Valencia Alfonso, Daelemans Walter, "Machine reading of biomedical texts about Alzheimers disease", CLEF 2012 Conference and Labs of the Evaluation Forum - Question Answering For Machine Reading Evaluation (QA4MRE), Rome / Forner, J. [edit.], e.a., 2012, p. 1-14
- Morante Roser, Sporleder Caroline, "Modality and negation: an introduction to the special issue", Florham Park, N.J., Computational linguistics, vol. 38, nr. 2, 2012, p. 223-260
- Oliveira Hugo Gonçalo, Burtenshaw Benjamin, Kestemont Mike, De Smedt Tom, "Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017)", Santiago de Compostela, Association for Computational Linguistics, 2017, 51 p.
- Ons Bart, Tessema Netsanet, van de Loo Janneke, Gemmeke Jort, De Pauw Guy, Daelemans Walter, Van hamme Hugo, "A self learning vocal interface for speech-impaired users", Stroudsburg, Pa, ACL, Proceedings 4th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT), 2013, p. 73-82
- Peersman Claudia, Daelemans Walter, Vandekerckhove Reinhild, Vandekerckhove Bram, van Vaerenbergh Leona, "The effects of age, gender and region on non-standard linguistic variation in online social networks", Antwerp, 2016, 24 p.
- Peersman Claudia, Vaassen Frederik, van Asch Vincent, Daelemans Walter, "Conversation level constraints on pedophile detection in chat rooms", Rome, CLEF 2012 Conference and Labs of the Evaluation Forum: Uncovering Plagiarism, Authorship, and Social Software Misuse (PAN), 2012, p. 1-13
- Peños Anselmo, Hovy Eduardo, Former Pamela, Rodrigo Alvaro, Morante Roser, "QA4MRE 2011-2013: overview of question answering for machine reading evaluation", Berlin, Springer, Lecture notes in computer science , 8138, 2013, p. 303-320
- Rangel Francisco, Rosso Paolo, Chugur Irina, Potthast Martin, Trenkmann Martin, Stein Benno, Verhoeven Ben, Daelemans Walter, "Overview of the 2nd Author Profiling Task at PAN 2014", CLEF 2014 Evaluation Labs and Workshop Working Notes Papers, Sheffield, UK, 2014, 2014, p. 1-30
- Rangel Francisco, Rosso Paolo, Verhoeven Ben, Daelemans Walter, Potthast Martin, Stein Benno, "Overview of the 4th Author Profiling Task at PAN 2016: Cross-genre Evaluations", Evora, CLEF and CEUR-WS.org., Working Notes Papers of the CLEF 2016 Evaluation Labs. CEUR Workshop Proceedings / Balog, Krisztian [edit.], et al., 2016, p. 750-784
- Rangel Pardo Francisco Manuel, Celli Fabio, Rosso Paolo, Potthast Martin, Stein Benno, Daelemans Walter, "Overview of the 3rd Author Profiling Task at PAN 2015", Toulouse, CEUR Workshop Proceedings , 1391, 2015, p. 1-8
- Rehm Georg, Uszkoreit Hans, Ananiadou Sophia, Bel Nuria, Bieleveciene Audrone, Borin Lars, Branco Antonio, Budin Gerhard, Calzolari Nicoletta, Daelemans Walter, Garabik Radovan, Grobelnik Marko, Garcia-Mateo Carmen, van Genabith Josef, Hajic Jan, Hernaez Inma, Judge John, "The strategic impact of META-NET on the regional, national and international level", New York, N.Y., Language resources and evaluation, vol. 50, nr. 2, 2016, p. 351-374
- Rehm Georg, Uszkoreit Hans, Ananiadou Sophia, Bel Nuria, Bieleveciene Audrone, Borin Lars, Branco Antonio, Budin Gerhard, Calzolari Nicoletta, Daelemans Walter, Garabik Radovan, Grobelnik Marko, Garcia-Mateo Carmen, van Genabith Josef, Hajic Jan, Hernaez Inma, Judge John, "The strategic impact of META-NET on the regional, national and international level", Paris, European language resources assoc-elra, LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 2014, p. 1517-1524
- Reviers Nina, Remael Aline, Daelemans Walter, "The language of audio description in Dutch: results of a corpus study", Bern, Peter Lang, New trends in translation studies , 15, 2015, p. 167-189
- Ruppenhofer J., Lee-Goldman R., Sporleder C., Morante Roser, "Beyond sentence-level semantic role labeling: linking argument structures in discourse", New York, N.Y., Language resources and evaluation, vol. 47, nr. 3, 2013, p. 695-721

- Rybicki Jan, Hoover David, Kestemont Mike, "Collaborative authorship: Conrad, Ford and Rolling Delta", Oxford, *Literary and linguistic computing*, vol. 29, nr. 3, 2014, p. 422-431
- Rys Kathy, Daelemans Walter, De Pauw Guy, Gillis Steven, "Leermechanismen in tweededialectverwerving", Groningen, *Tabu: bulletin voor taalwetenschap*, vol. 40, nr. 2012, p. 59-88
- Rys Kathy, Keuleers Emmanuel, Daelemans Walter, Gillis Steven, "Acquisition of phonological variables of a Flemish dialect by children raised in Standard Dutch: some considerations on the learning mechanisms", Amsterdam, *Benjamins, Studies in Language Variation*, 20, 2017, p. 267-304
- Sabbe Bernard, Beheydt Lieve, De Picker Livia, Goetschalckx Jens, Daelemans Walter, "Computational language analysis for assessment of schizophrenia", *Proceedings 6th International Conference on Cognitive and Behavioral Psychology (CBP)*, 2017, p. 118-127
- Sandra Dominiek, "Mentale processen achter d/t-fouten: is een verklaring ook een legitimatie?", Kortrijk-Heule, UGA, *Over taal: tijdschrift over taal, tekst en communicatie*, vol. 54, 2015, p. 34-37
- Scheurwegs Alyne, Luyckx Kim, Luyten Leon, Daelemans Walter, van den Bulcke Tim, "Data integration of structured and unstructured sources for assigning clinical codes to patient stays", Philadelphia, Pa, *Journal of the American Medical Informatics Association*, vol. 23, 2016, p. 11-19
- Scheurwegs Elyne, Cule Boris, Luyckx Kim, Luyten Leon, Daelemans Walter, "Selecting relevant features from the electronic health record for clinical code prediction", New York, *Journal of biomedical informatics*, nr. 74, 2017, p. 92-103
- Scheurwegs Elyne, Luyckx Kim, Luyten Leon, Goethals Bart, Daelemans Walter, "Assigning clinical codes with data-driven concept representation on Dutch clinical free text", New York, *Journal of biomedical informatics*, vol. 69, 2017, p. 118-127
- Scheurwegs Elyne, Luyckx Kim, van der Schueren Filip, van den Bulcke Tim, "De-identification of clinical free text in Dutch with limited training data: a case study", S.l., *Workshop on NLP for Medicine and Biology*, 2013, p. 18-23
- Scheurwegs Elyne, Sushil Madhumita, Tulkens Stéphan, Daelemans Walter, Luyckx Kim, "Counting trees in Random Forests: predicting symptom severity in psychiatric intake reports", New York, *Journal of biomedical informatics*, vol. 75, Supplement, 2017, p. 112-119
- Schuerewegen Franc, "Brèves remarques sur Franco Moretti", Paris, *Seuil, Poétique: collection*, nr. 169, 2012, p. 89-96
- Schulz Sarah, De Pauw Guy, De Clercq Orphee, Desmet Bart, Hoste Veronique, Daelemans Walter, Macken Lieve, "Multimodular text normalization of Dutch user-generated content", *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 7, nr. 4, 2016, p. 1-22
- Snyman Dirk, van Huyssteen Gerhard, Daelemans Walter, "Cross-lingual genre classification for closely related languages", Pretoria, *Proceedings of the Twenty-Third Annual Symposium of the Pattern Recognition Association of South Africa*, 2012, p. 133-137
- Snyman Dirk, van Huyssteen Gerhard, Daelemans Walter, "Outomatiese genreklassifikasie vir Afrikaans", Pretoria, *Die Suid-Afrikaanse tydskrif vir natuurwetenskap en tegnologie*, vol. 33, nr. 1, 2014, p. 1-12
- Stamatatos Efstathios, Daelemans Walter, Verhoeven Ben, Juola Patrick, López-López Aurelio, Potthast Martin, Stein Benno, "Overview of the author identification task at PAN 2015", Toulouse, *CEUR workshop proceedings*, 1391, 2015, p. 1-17
- Stamatatos Efstathios, Daelemans Walter, Verhoeven Ben, Potthast Martin, Stein Benno, Juola Patrick, Sanchez-Perez Miguel A., Barron-Cedeno Alberto, "Overview of the Author Identification Task at PAN 2014", *CLEF 2014 Evaluation Labs and Workshop Working Notes Papers*, Sheffield, UK, 2014, 2014, p. 1-21
- Stamatatos Estathios, Tschuggnall Michael, Verhoeven Ben, Daelemans Walter, Specht Gunther, Stein Benno, Potthast Michael, "Clustering by authorship within and across documents", Evora, *CLEF and CEUR-WS.org., Working Notes Papers of the CLEF 2016 Evaluation Labs. CEUR Workshop Proceedings / Balog, Krisztian [edit.]*, et al., 2016, p. 691-715
- Stover J.A., Kestemont Mike, "Reassessing the Apuleian Corpus: a computational approach to authenticity", Oxford, *The classical quarterly*, 2016, p. 1-44
- Stover Justin A., Kestemont Mike, "The authorship of the 'Historia Augusta': two new computational studies", London, *Bulletin of the Institute of Classical Studies of the University of London*, vol. 59, nr. 2, 2016, p. 140-157
- Stover Justin Anthony, Winter Yaron, Koppel Moshe, Kestemont Mike, "Computational authorship verification method attributes a new work to a major 2nd century African author", New York, N.Y., Wiley, *Journal of the Association for Information Science and Technology*, vol. 67, nr. 1, 2016, p. 239-242
- Sushil Madhumita, Suster Simon, Luyckx Kim, Daelemans Walter, "Unsupervised patient representations from clinical notes with interpretable classification decisions", 2017, 6 p.

- Suster Simon, van Noord Gertjan, Titov Ivan, "Word representations: tree models and syntactic functions", 2015, 11 p.
- Suster Simon, "Empirical studies on word representations", Groningen, Rijksuniversiteit Groningen, 2016, 232 p.
- Suster Simon, Titov Ivan, van Noord Gertjan, "Bilingual learning of multi-sense embeddings with discrete autoencoders", San Diego, CA, Association for Computational Linguistics, Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, p. 1346-1356
- Suster Simon, Tulkens Stéphan, Daelemans Walter, "A short review of ethical challenges in clinical natural language processing", Valencia, Association for Computational Linguistics, Proceedings of the First ACL Workshop on Ethics in Natural Language Processing / Hovy, Dirk [edit.], et al., 2017, p. 80-87
- Tessema Netsanet, Ons Bart, van de Loo Janneke, Gemmeke Jort, De Pauw Guy, Daelemans Walter, "Metadata for Corpora PATCOR and Domotica-2", Leuven, KUL, Technical report, KUL/ESAT/PSI/1303, 2013, 5 p.
- Tobback Ellen, Naudts Hans, Daelemans Walter, Junqué de Fortuny Enric, Martens David, "Belgian economic policy uncertainty index: improvement through text mining", Amsterdam, International journal of forecasting, 2016, p. 1-11
- Tschuggnall Michael, Stamatatos Efsthios, Verhoeven Ben, Daelemans Walter, Specht Günther, Stein Benno, Potthast Martin, "Overview of the Author Identification Task at PAN-2017: Style Breach Detection and Author Clustering", CEUR Workshop Proceedings, 1866, 2017, p. 1-22
- Tulkens Stéphan, Emmerly Chris, Daelemans Walter, "Evaluating unsupervised Dutch word embeddings as a linguistic resource", Portorož, European Language Resources Association (ELRA), Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016) / Calzolari, Nicoletta [edit.], et al., 2016, p. 1-7
- Tulkens Stéphan, Hilde Lisa, Lodewyckx Elise, Verhoeven Ben, Daelemans Walter, "A dictionary-based approach to racism detection in Dutch social media", Portorož, European Language Resources Association (ELRA), Proceedings of the first Workshop on Text Analytics for Cybersecurity and Online Safety (TA-COS 2016) / Daelemans, Walter [edit.], et al., 2016, p. 1-7
- Tulkens Stéphan, Hilde Lisa, Lodewyckx Elise, Verhoeven Ben, Daelemans Walter, "The automated detection of racist discourse in Dutch social media", Computational Linguistics in the Netherlands Journal, vol. 6, nr. 1, 2016, p. 3-20
- Tulkens Stéphan, Suster Simon, Daelemans Walter, "Using distributed representations to disambiguate biomedical and clinical concepts", Berlin, Association for Computational Linguistics, Proceedings of the 15th Workshop on Biomedical Natural Language Processing of the Association for Computational Linguistics, 2016, p. 77-82
- Vaassen Frederik, "Measuring emotion: exploring the feasibility of automatically classifying emotional tekst", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departement Taalkunde, 2014, 127 p.
- Vaassen Frederik, Wauters Jeroen, Van Broekhoven Frederik, Van Overveldt Maarten, Eneman Koen, Daelemans Walter, "deLearyous: training interpersonal communication skills using unconstrained text input", Reading, Academic Publishing International Limited, Proceedings of ECGBL 2012, The 6th European Conference on Games Based Learning / Felicia, Patrick [edit.], 2012, p. 505-513
- van Asch Vincent, "Domain similarity measures: on the use of distance metrics in natural language processing", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departement Taalkunde, 2012, 200 p.
- van Asch Vincent, Daelemans Walter, "An analytical approach to similarity measure selection for selftraining", BENELEARN 2013: proceedings of the 22nd Belgian-Dutch Conference on Machine Learning, Nijmegen, the Netherlands, 2013, p. 8-17
- van de Loo Janneke, "Machine learning approaches to modeling language use of individuals and demographic groups: applications in societally relevant natural language processing", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departement Taalkunde, 2016, 126 p.
- van de Loo Janneke, De Pauw Guy, Gemmeke Jort, Karsmakers Peter, Van Den Broeck Bert, Daelemans Walter, Van hamme Hugo, "Towards shallow grammar induction for an adaptive assistive vocal interface: a concept tagging approach", Istanbul, Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility, 2012, p. 27-34
- van de Loo Janneke, Gemmeke Jort F., De Pauw Guy, Driesen Joris, Van Hamme Hugo, Daelemans Walter, "Towards a self-learning assistive vocal interface: vocabulary and grammar learning", Jeju, Proceedings of the 1st Workshop on Speech and Multimodal Interaction in Assistive Environments, 2012, p. 34-42

- van den Bosch Antal, Daelemans Walter, "Implicit schemata and categories in memory-based language processing", Teddington, Md, Language and speech, vol. 56, nr. 3, 2013, p. 309-328
- Van Hee Cynthia, Lefever Els, Verhoeven Ben, Mennes Julie, Desmet Bart, De Pauw Guy, Daelemans Walter, Hoste Véronique, "Automatic detection and prevention of cyberbullying", St. Julians, IARIA, Proceedings of the First International Conference on Human and Social Analytics (HUSO 2015) / Lorenz, Pascal [edit.], et al., 2015, p. 13-18
- Van Hee Cynthia, Lefever Els, Verhoeven Ben, Mennes Julie, Desmet Bart, De Pauw Guy, Daelemans Walter, Hoste Veronique, "Detection and fine-grained classification of cyberbullying events", Proceedings of the 10th Recent Advances in Natural Language Processing (RANLP 2015), 2015, p. 672-680
- Van Hee Cynthia, Verhoeven Ben, Lefever Els, De Pauw Guy, Hoste Véronique, Daelemans Walter, "Guidelines for the fine-grained analysis of cyberbullying", LT3 Technical report series, 15-01, 2015, 16p.
- Van Hulle Dirk, "A James Joyce digital library", Amsterdam, European Joyce studies, vol. 25, 2016, p. 226-246
- Van Hulle Dirk, "Beckett's negative capability and digital genetic editing: decision-making and creative hesitation", Vinhedo, Editora Horizonte, Estudos crítica genética, 2016, p. 189-198
- Van Hulle Dirk, "De autograaf en het pentimenti-model: een verkenning van het Claus-archief", Antwerpen, Spiegel der letteren, vol. 58, nr. 2, 2016, p. 157-171
- Van Hulle Dirk, "Digital library history: the virtual bookcases of James Joyce and Samuel Beckett", Amsterdam, Quaerendo: a quarterly journal from the Low Countries devoted to manuscripts and printed books, vol. 46, nr. 2-3, 2016, p. 192-204
- Van Hulle Dirk, "Exogenetic digital editing and enactive cognition", Open Book Publishers, Digital humanities series, 4, 2016, p. 107-115
- Van Hulle Dirk, "Flann O'Brien's Ulysses: marginalia and the modernist mind", London, Bloomsbury, Flann O'Brien and modernism, 2014, p. 107-119
- Van Hulle Dirk, "La genèse de L'Innommable: essai de critique génétique postcognitiviste", Paris, Littérature, nr. 167, 2012, p. 65-77
- Van Hulle Dirk, "Modelling a digital scholarly edition for genetic criticism: a rapprochement", Turnhout, Variants: the journal of the European Society for Textual Scholarship, vol. 12-13, 2016, p. 34-56
- Van Hulle Dirk, "Modern manuscripts: the extended mind and creative undoing from Darwin to Beckett and beyond", London, Bloomsbury, 2014
- Van Hulle Dirk, "Textual scars: Beckett, genetic criticism and textual scholarship", Edinburgh, Edinburgh University Press, The Edinburgh companion to Samuel Beckett and the arts, 2014, p. 306-319
- Van Hulle Dirk, "The Obidil and the Man of glass: denarration, genesis and cognition in Beckett's Molloy, Malone dies and The unnamable", Amsterdam, Rodopi, Samuel Beckett today: an annual bilingual review, vol. 26, 2014, p. 25-40
- Van Hulle Dirk, "The stuff of fiction: digital editing, multiple drafts and the extended mind", Bloomington, Ind., Indiana University Press, Textual cultures: texts, contexts, interpretation, vol. 8, nr. 1, 2013, p. 23-37
- Van Hulle Dirk, Becketts Mancando, Schoten, "Deus ex machina: tijdschrift voor actuele literatuur", vol. 38, nr. 150, 2014, p. 42-47
- Van Hulle Dirk, Kestemont Mike, "Periodizing Samuel Beckett's Works A Stylochronometric Approach", Fayetteville, Ark., Style, vol. 50, nr. 2, 2016, p. 172-202
- Van Hulle Dirk, Kestemont Mike, "Stylochronometry and the Periodization of Samuel Beckett", Kraków, Proceedings of Digital Humanities 2016, 2016, p. 393-395
- Van Royen Kathleen, Poels Karolien, Daelemans Walter, Vandebosch Heidi, "Automatic monitoring of cyberbullying on social networking sites: from technical feasibility to desirability", New York, Telematics and informatics, vol. 32, nr. 1, 2015, p. 89-97
- van Zaanen Menno, van de Loo Janneke, "Learning interpretations using sequence classification", JMLR: workshop and conference proceedings, 21, 2013, p. 220-223
- van Zundert Joris, Chambers Sally, Kestemont Mike, Koolen Marijn, Jones Catherine, "Proceedings of the DHBeNeLux Conference 2015", DHQ, vol. 11, nr. 4, 2017
- Vandekerckhove Bram, "Exemplar-based generalization at the interface between syntax and semantics", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departement Taalkunde, 2013, 172 p.
- Vandekerckhove Bram, Sandra Dominiek, Daelemans Walter, "Effects of online abstraction on adjective order preferences", Abingdon, Language, cognition and neuroscience, vol. 30, nr. 7, 2015, p. 816-831,
- Vandekerckhove Bram, Sandra Dominiek, Daelemans Walter, "Selective impairment of adjective order constraints as overeager abstraction: an elaboration on Kemmerer et al. (2009)", Oxford, Journal of neurolinguistics, vol. 26, nr. 1, 2013, p. 46-72

- Vanormelingen Liesbeth, Gillis Steven, The influence of socio-economic status on mothers' volubility and responsiveness in a monolingual Dutch-speaking sample, Chalfont St. Giles, First language, vol. 36, nr. 2, 2016, p. 140-156
- Verbeke Mathias, Frasconi Paolo, van Asch Vincent, Morante Roser, Daelemans Walter, De Raedt Luc, "Kernel-based logical and relational learning with kLog for hedge cue detection", Berlin, Springer, Inductive Logic Programming / Muggleton, Stephen [edit.], e.a., 2012, p. 347-357
- Verbeke, M., Frasconi, P., De Grave, K., Costa, F., De Raedt, L. (2014). kLogNLP: Graph kernel-based relational learning of natural language. In Bontcheva, K. (Ed.), Jingbo, Z. (Ed.), Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Annual Meeting of the Association for Computational Linguistics (ACL). Baltimore, Maryland, USA, 22-27 June 2014 (pp. 85-90) Association for Computational Linguistics.
- Verbeke Mathias, Morante Roser, van Asch Vincent, Frasconi Paolo, Daelemans Walter, De Raedt Luc, "A statistical relational learning approach to identifying evidence based medicine categories", Association for Computational Linguistics, Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 2012, p. 579-589
- Verbeke Mathias, van Asch Vincent, Daelemans Walter, De Raedt Luc, "Lazy and eager relational learning using Graph-Kernels", LNAI/LNCS proceedings (Springer) , 8791, 2014, p. 163-176
- Verhaert Nina, "Rules or regularities? The homephone dominance effect in spelling and reading regular Dutch verb forms", Antwerpen, Universiteit Antwerpen, Faculteit Letteren en Wijsbegeerte, Departement Taalkunde, 2016, 457 p.
- Verhoeven Ben, van Huyssteen Gerhard B., "More than only noun-noun compounds: towards an annotation scheme for the semantic modelling of other noun compound types", Potsdam, Association for Computational Linguistics, Proceedings of the 9th Joint ISO, ACL SIGSEM Workshop on Interoperable Semantic Annotation, 2013, p. 59-66
- Verhoeven Ben, van Huyssteen Gerhard, van Zaanen Menno, Daelemans Walter, "Annotation guidelines for compound analysis", CLiPS technical report series, nr. 5, 2014, p. 1-50
- Verhoeven Ben, van Zaanen Menno, Daelemans Walter, van Huyssteen Gerhard, "Automatic compound processing: compound splitting and semantic analysis for Afrikaans and Dutch", Proceedings of the First Workshop on Computational Approaches to Compound Analysis, (ComAComA 2014), Dublin, Ireland, August 24, 2014 / Verhoeven, B. [edit.], e.a., 2014, p. 20-30
- Verhoeven Ben, Daelemans Walter, van Huyssteen Gerhard, "Classification of noun-noun compound semantics in Dutch and Afrikaans", Pretoria, Proceedings of the Twenty-Third Annual Symposium of the Pattern Recognition Association of South Africa, 2012, p. 121-125
- Verhoeven Ben, Daelemans Walter, van Zaanen Menno, van Huyssteen Gerhard, "Proceedings of the First Workshop on Computational Approaches to Compound Analysis held at the 25th International Conference on Computational Linguistics (COLING 2014)", Dublin, 2014, 91 p.
- Verhoeven Ben, Daelemans Walter, "CLiPS Stylometry Investigation (CSI) corpus: a Dutch corpus for the detection of age, gender, personality, sentiment and deception in text", Paris, European language resources assoc-elra, LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 2014, p. 3081-3085
- Verhoeven Ben, Daelemans Walter, "Semantic classification of Dutch noun-noun compounds: a distributional semantics approach", Computational linguistics in the Netherlands journal, nr. 3, 2013, p. 2-18
- Verhoeven Ben, Daelemans Walter, De Pauw Guy, Kestemont Mike, Preface, Computational linguistics in the Netherlands journal, nr. 5, 2015, p. 1-2
- Verhoeven Ben, Daelemans Walter, De Smedt Tom, "Ensemble methods for personality recognition", Proceedings of the Workshop on Computational Personality Recognition (Shared Task), Cambridge, Mass., 2013, p. 1-4
- Verhoeven Ben, Daelemans Walter, Plank Barbara, "Creating TwiSty: corpus development and statistics", CLiPS Technical Report Series (CTRS), nr. 6, 2016, p. 1-21
- Verhoeven Ben, Daelemans Walter, Plank Barbara, "TwiSty: a multilingual Twitter Stylometry corpus for gender and personality profiling", Portorož , European Language Resources Association (ELRA), Proceedings of the 10th Annual Conference on Language Resources and Evaluation (LREC 2016) / Calzolari, Nicoletta [edit.], et al., 2016, p. 1-6
- Verhoeven Ben, Soler Juan, Daelemans Walter, "Evaluating content-independent features for personality recognition", New York, ACM, WCPR '14 Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition, 2014, p. 7-10

Wauters Jeroen, Van Broeckhoven Frederik, Van Overveldt Maarten, Eneman Koen, Vaassen Frederik, Daelemans Walter, "deLearyous: an interactive application for interpersonal communication training, Berlin", Springer, Serious games: the challenge / Wannemacker, De, Stefan [edit.], e.a., 2012, p. 87-90

C. UNIVERSITEIT GENT - OVERZICHT PUBLICATIES

- Amelynck, Denis, Pieter Maes, Marc Leman, et al. "The Surprising Character of Music: a Search for Sparsity in Music Evoked Body Movements." *Analysis of Large and Complex Data*. Ed. Adalbert Wilhelm & Hans Kestler. Springer International Publishing, 2016. 425–435.
- Amelynck, Denis, Pieter-Jan Maes, Jean-Pierre Martens, et al. "Beating-time Gestures Imitation Learning for Humanoid Robots." *EAI ENDORSED TRANSACTIONS ON CREATIVE TECHNOLOGIES* 4.13 (2017): n. pag.
- Amelynck, Denis, Pieter-Jan Maes, Jean-Pierre Martens, et al. "Expressive Body Movement Responses to Music Are Coherent, Consistent, and Low Dimensional." *IEEE TRANSACTIONS ON CYBERNETICS* 44.12 (2014): 2288–2301.
- Bordea, Georgeta, Els Lefever, and Paul Buitelaar. "SemEval-2016 Task 13: Taxonomy Extraction Evaluation (TExEval-2)." *Proceedings of SemEval-2016*. San Diego, California: Association for Computational Linguistics, 2016. 1081–1091.
- Clapham, Renee, Catherine Middag, Frans Hilgers, et al. "Developing Automatic Articulation, Phonation and Accent Assessment Techniques for Speakers Treated for Advanced Head and Neck Cancer." *SPEECH COMMUNICATION* 59 (2014): 44–54.
- Clapham, Renee, Jean-Pierre Martens, Rob Van Son, et al. "Computing Scores of Voice Quality and Speech Intelligibility in Tracheoesophageal Speech for Speech Stimuli of Varying Lengths ." *COMPUTER SPEECH AND LANGUAGE* 37 (2016): 1–10.
- Daems, Joke, and Lieve Macken. "Faster Is Not Always Better, but What If It's Not Worse? A Comparative Study of Human Translation and Post-edited Machine Translation Quality." 2013 : n. pag.
- Daems, Joke, Lennart Tondeleir, and Lieve Macken. "Machine Translation: Friend or Foe? Translator Attitude, Process, Productivity and Quality in Human Translation and Post-editing Machine Translation." 2013 : n. pag.
- Daems, Joke, Lieve Macken, and Sonia Vandepitte. "On the Origin of Errors: a Fine-grained Analysis of MT and PE Errors and Their Relationship." *LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION*. Ed. Nicoletta Calzolari et al. European Language Resources Association (ELRA), 2014. 62–66.
- Daems, Joke, Lieve Macken, and Sonia Vandepitte. "Quality as the Sum of Its Parts: a Two-step Approach for the Identification of Translation Problems and Translation Quality Assessment for HT and MT+PE." *MT Summit XIV Workshop on Post-editing Technology and Practice, Proceedings*. Ed. Sharon O'Brien, Michel Simard, & Lucia Specia. European Association for Machine Translation, 2013. 63–71.
- Daems, Joke, Lieve Macken, and Sonia Vandepitte. "Snelheid Vs. Kwaliteit: Schuilt Er Wel Winst in Het Gebruik Van Automatische Vertaalsystemen?" *Beschouwingen Uit Een Talenhuis : Opstellen over Onderwijs En Onderzoek in De Vakgroep Vertalen, Tolken En Communicatie Aangeboden Aan Rita Godyns*. Ed. Stefaan Evenepoel, Patrick Goethals, & Lieve Jooken. Gent: Academia Press, 2014. 141–146.
- Daems, Joke, Lieve Macken, and Sonia Vandepitte. "Two Sides of the Same Coin: Assessing Translation Quality in Two Steps Through Adequacy and Acceptability Error Analysis." *LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION*. 2014. 2–2.
- Daems, Joke, Lieve Macken, Sonia Vandepitte, et al. "Lost in Post-editing: Identifying the MT Error Types That Are Most Problematic for Student Post-editors." 2014 : n. pag.
- Daems, Joke, Michael Carl, Sonia Vandepitte, et al. "How Do Students Cope with Machine Translation Output of Multi-Word Units? An Exploratory Study." *Multi-word Units in Machine Translation and Translation Technology*. Ed. Ruslan Mitkov et al. John Benjamins, 2016.
- Daems, Joke, Michael Carl, Sonia Vandepitte, et al. "The Effectiveness of Consulting External Resources During Translation and Postediting of General Text Types." *New Directions in Empirical Translation Process Research : Exploring the CRITT TPR-DB*. Ed. Michael Carl, Srinivas Bangalore, & Moritz Schaeffer. Springer, 2016. 111–133.
- Daems, Joke, Orphée De Clercq, and Lieve Macken. "Translationese and Post-editese : How Comparable Is Comparable Quality?" *LINGUISTICA ANTVERPIENSIA NEW SERIES-THEMES IN TRANSLATION STUDIES* (2017): n. pag.
- Daems, Joke, Sonia Vandepitte, Robert Hartsuiker, et al. "Comparing HT and PE Using Advanced Research Tools." 2015 : n. pag.
- Daems, Joke, Sonia Vandepitte, Robert Hartsuiker, et al. "The Impact of Machine Translation Error Types on Post-editing Effort Indicators." *Fourth Workshop on Post-Editing Technology and Practice, Proceedings*. Ed. Sharon O'Brien & Michel Simard. Association for Machine Translation in the Americas, 2015. 31–45.

- Daems, Joke, Sonia Vandepitte, Robert Hartsuiker, et al. "Translation Methods and Experience : a Comparative Analysis of Human Translation and Post-editing with Students and Professional Translators." *META* 62.2 (2017): 245–270.
- Daems, Joke, Sonia Vandepitte, Robert J. Hartsuiker, et al. "Identifying the Machine Translation Error Types with the Greatest Impact on Post-editing Effort." Ed. Anna Esposito. *FRONTIERS IN PSYCHOLOGY* 8 (2017): n. pag.
- De Clercq, Orphée, and Veronique Hoste. "All Mixed up? Finding the Optimal Feature Set for General Readability Prediction and Its Application to English and Dutch." *COMPUTATIONAL LINGUISTICS* 42.3 (2016): 457–490.
- De Clercq, Orphée, and Veronique Hoste. "Comparing Features for Automatic Readability Prediction." *ATILA Workshop*. 2012.
- De Clercq, Orphée, and Veronique Hoste. "Crash-testing Text Processing on User-generated Content." *Computational Linguistics in the Netherlands*. 2014.
- De Clercq, Orphée, and Veronique Hoste. "Features for Automatic Readability Prediction." *Computational Linguistics in the Netherlands*. 2013.
- De Clercq, Orphée, and Veronique Hoste. "Hoe Meetbaar Is Leesbaarheid?" *Beschouwingen Uit Een Talenhuis : Opstellen over Onderwijs En Onderzoek in De Vakgroep Vertalen, Tolken En Communicatie Aangeboden Aan Rita Godyns*. Ed. Stefaan Evenepoel, Patrick Goethals, & Lieve Jooken. Ghent, Belgium: Academia Press, 2014. 147–155.
- De Clercq, Orphée, and Veronique Hoste. "Plain Language: Investigating the Effect of Different Text Characteristics on Text Readability." *3rd Translation Studies Day*. 2013.
- De Clercq, Orphée, and Veronique Hoste. "Rude Waiter but Mouthwatering Pastries! An Exploratory Study into Dutch Aspect-based Sentiment Analysis." *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Ed. Nicoletta Calzolari et al. Paris, France: ELRA, 2016. 2910–2917.
- De Clercq, Orphée, Bart Desmet, Sarah Schulz, et al. "Normalization of Dutch User-generated Content." *Proceedings of Recent Advances in Natural Language Processing*. Shoumen, Bulgaria: INCOMA, 2013. 179–188.
- De Clercq, Orphée, Els Lefever, Gilles Jacobs, et al. "Hello ABSA! Collaboration Between LT3 and Hello Customer to Develop an Aspect-based Sentiment Analysis Pipeline." 2017.
- De Clercq, Orphée, Els Lefever, Gilles Jacobs, et al. "Towards an Integrated Pipeline for Aspect-based Sentiment Analysis in Various Domains." *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Association for Computational Linguistics, 2017. 136–142.
- De Clercq, Orphée, Marjan Van de Kauter, Els Lefever, et al. "Applying Hybrid Terminology Extraction to Aspect-based Sentiment Analysis." *International Workshop on Semantic Evaluation, Proceedings*. Association for Computational Linguistics, 2015. 719–724.
- De Clercq, Orphée, Marjan Van de Kauter, Els Lefever, et al. "Terminology Extraction to Aspect-Based Sentiment Analysis." *Computational Linguistics in the Netherlands*. 2015.
- De Clercq, Orphée, Michael Schuhmacher, Simone Paolo Ponzetto, et al. "Content-Based Book Recommendation." *ATILA Workshop*. 2014.
- De Clercq, Orphée, Michael Schuhmacher, Simone Paolo Ponzetto, et al. "Exploiting frameNet for Content-based Book Recommendation." *CBRecSys at ACM RecSys, Proceedings*. CEUR-WS, 2014. 14–21.
- De Clercq, Orphée, Sarah Schulz, Bart Desmet, et al. "Normalization of Dutch User-Generated Content." *Computational Linguistics in the Netherlands*. 2014.
- De Clercq, Orphée, Sarah Schulz, Bart Desmet, et al. "Towards Shared Datasets for Normalization Research." *LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION*. Ed. Nicoletta Calzolari et al. European Language Resources Association (ELRA), 2014. 1218–1223.
- De Clercq, Orphée, Sarah Schulz, Els Lefever, et al. "First Steps in Normalizing Dutch UGC." *ATILA Workshop*. 2013.
- De Clercq, Orphée, Sven Hertling, Veronique Hoste, et al. "Identifying Disputed Topics in the News." *Proceedings of the LD4KD Workshop at ECML/PKDD2014*. CEUR, 2014. 37–48.
- De Clercq, Orphée, Veronique Hoste, and Paola Monachesi. "Evaluating Automatic Cross-domain Dutch Semantic Role Annotation." *Computational Linguistics in the Netherlands*. 2012.
- De Clercq, Orphée, Veronique Hoste, and Paola Monachesi. "Evaluating Automatic Cross-domain Dutch Semantic Role Annotation." *LREC 2012 : Eight International Conference on Language Resources and Evaluation*. Ed. Nicoletta Calzolari et al. Paris, France: European Language Resources Association (ELRA), 2012. 88–93.
- De Clercq, Orphée, Veronique Hoste, Bart Desmet, et al. "Using the Crowd for Readability Prediction." *NATURAL LANGUAGE ENGINEERING* 20.3 (2014): 293–325.
- Dejonckere, Philippe, KJ Neumann, Mieke Moerman, et al. "Tridimensional Assessment of Adductor Spasmodic Dysphonia Pre- and Post-treatment with Botulinum Toxin." *EUROPEAN ARCHIVES OF OTO-RHINO-LARYNGOLOGY* 269.4 (2012): 1195–1203.

- Dejonckere, Philippe, Mieke Moerman, Jean-Pierre Martens, et al. "Voicing Quantification Is More Relevant Than Period Perturbation in Substitution Voices: An Advanced Acoustical Study." *EUROPEAN ARCHIVES OF OTO-RHINO-LARYNGOLOGY* 269.4 (2012): 1205–1212.
- Demuynck, Kris, and Fabian Triefenbach. "Porting Concepts from DNNs Back to GMMs." 2013 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU). New York, NY, USA: IEEE, 2013. 356–361.
- Desmet, Bart, and Veronique Hoste. "Combining Lexico-semantic Features for Emotion Classification in Suicide Notes." *BIOMEDICAL INFORMATICS INSIGHTS* 5.suppl. 1 (2012): 125–128.
- Desmet, Bart, and Veronique Hoste. "Emotion Detection in Suicide Notes." *EXPERT SYSTEMS WITH APPLICATIONS* 40.16 (2013): 6351–6358.
- Desmet, Bart, and Veronique Hoste. "Final Words: Topic Detection in Suicide Notes." 2012.
- Desmet, Bart, and Veronique Hoste. "Fine-grained Dutch Named Entity Recognition." *LANGUAGE RESOURCES AND EVALUATION* 48.2 (2014): 307–343.
- Desmet, Bart, and Veronique Hoste. "Recognising Suicidal Messages in Dutch Social Media." *LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION*. 2014. 830–835.
- Desmet, Bart, and Veronique Hoste. "Suicidality Detection in Social Media." 2015.
- Desmet, Bart, Gilles Jacobs, and Veronique Hoste. "Mental Distress Detection and Triage in Forum Posts: The LT3 CLPsych 2016 Shared Task System." *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*. Association for Computational Linguistics, 2016. 148–152.
- Desmet, Bart, Kirsten Pauwels, and Veronique Hoste. "Online Suicide Risk Detection Using Automatic Text Classification." *International Summit on Suicide Research*. 2015.
- Desmet, Bart, Orphée De Clercq, Marjan Van de Kauter, et al. "Taaltechnologie 2.0: Sentimentanalyse En Normalisatie." *Beschouwingen Uit Een Talenhuis : Opstellen over Onderwijs En Onderzoek in De Vakgroep Vertalen, Tolken En Communicatie Aangeboden Aan Rita Godyns*. Ed. Stefaan Evenepoel, Patrick Goethals, & Lieve Jooken. Ghent, Belgium: Academia Press, 2014. 157–161.
- Desmet, Bart, Sander Naert, Veerle Ongenaes, et al. "Extending Brat, an Online Annotation Tool." *ATILA Workshop*. 2013.
- Desmet, Frank, Luc Nijs, Michiel Demey, et al. "Assessing a Clarinet Player's Performer Gestures in Relation to Locally Intended Musical Targets." *JOURNAL OF NEW MUSIC RESEARCH* 41.1 (2012): 31–48.
- Desplanques, Brecht, and Jean-Pierre Martens. "Model-based Speech/non-speech Segmentation of a Heterogeneous Multilingual TV Broadcast Collection." *International Symposium on Intelligent Signal Processing and Communication Systems ISPACS*. IEEE, 2013. 55–60.
- Desplanques, Brecht, Kris Demuynck, and Jean-Pierre Martens. "Adaptive Speaker Diarization of Broadcast News Based on Factor Analysis." *COMPUTER SPEECH AND LANGUAGE* 46 (2017): 72–93.
- Desplanques, Brecht, Kris Demuynck, and Jean-Pierre Martens. "Combining Joint Factor Analysis and iVectors for Robust Language Recognition." *Odyssey 2014 : The Speaker and Language Recognition Workshop*, Proceedings. 2014. 73–80.
- Desplanques, Brecht, Kris Demuynck, and Jean-Pierre Martens. "Factor Analysis for Speaker Segmentation and Improved Speaker Diarization." *16th Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)*, Vols 1-5. Baixas, France: International Speech Communication Association (ISCA), 2015. 3081–3085.
- Desplanques, Brecht, Kris Demuynck, and Jean-Pierre Martens. "Robust Language Recognition via Adaptive Language Factor Extraction." *15th Annual Conference of the International Speech Communication Association*, Proceedings. 2014. 2160–2164.
- Desplanques, Brecht, Kris Demuynck, and Jean-Pierre Martens. "Soft VAD in Factor Analysis Based Speaker Segmentation of Broadcast News." 2016. 158–165.
- Farasyn, Melissa, Mariya Koleva, Veronique Hoste, et al. "A Parsed Corpus of Historical Low German." Antwerp, Belgium, 2015.
- Farasyn, Melissa, Mariya Koleva, Veronique Hoste, et al. "Towards a Part-of-speech Tagger for Middle Low German." Groeningen, Netherlands, 2015.
- Gemmeke, Jort F, Tuomas Virtanen, and Kris Demuynck. "Exemplar-based Joint Channel and Noise Compensation." *International Conference on Acoustics Speech and Signal Processing ICASSP*. IEEE, 2013. 868–872.
- Goethals, Patrick, Els Lefever, and Lieve Macken. "SCAP_TT : Tagging and Lemmatising Spanish Tourism Discourse, and Beyond." *IBERICA* 33 (2017): 273–282.
- Haderlein, Tino, Catherine Middag, Andreas Maier, et al. "Visualization of Intelligibility Measured by Language-independent Features." Springer, 2014. 547–554.

- Haderlein, Tino, Catherine Middag, Florian Höning, et al. "Language-independent Age Estimation from Speech Using Phonological and Phonemic Features." *Lecture Notes in Artificial Intelligence*. Ed. P Král & V Matoušek. Vol. 9302. Switzerland: Springer International Publishing, 2015. 165–173.
- Haderlein, Tino, Catherine Middag, Jean-Pierre Martens, et al. "Language-independent Automatic Evaluation of Intelligibility of Chronically Hoarse Persons." *FOLIA PHONIATRICA ET LOGOPAEDICA* 66.6 (2014): 219–226.
- Hanouille, Sabien, Veronique Hoste, and Aline Remael. "The Efficacy of Terminology-extraction Systems for the Translation of Documentaries." *PERSPECTIVES-STUDIES IN TRANSLATOLOGY* 23.3 (2015): 359–374.
- Hanouille, Sabien, Veronique Hoste, and Aline Remael. "The Translation of Documentaries: Can Domain-specific, Bilingual Glossaries Reduce the Translators' Workload? An Experiment Involving Professional Translators." *NEW VOICES IN TRANSLATION STUDIES* 23 (2015): 25–49.
- Hendrickx, Iris, Els Lefever, Ilja Croijmans, et al. "Very Quaffable and Great Fun: Applying NLP to Wine Reviews." *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*. Berlin, Germany: Association for Computational Linguistics, 2016. 306–312.
- Hendrickx, Iris, Gosse Bouma, Walter Daelemans, et al. "COREA: Coreference Resolution for Extracting Answers for Dutch." *Essential Speech and Language Technology for Dutch : Results by the STEVIN-programme*. Berlin, Germany: Springer, 2013. 97–108.
- Hoste, Veronique, Cynthia Van Hee, and Karolien Poels. "Towards a Framework for the Automatic Detection of Crisis Emotions on Social Media : a Corpus Analysis of the Tweets Posted After the Crash of Germanwings Flight 9525." *Proceedings of HUSO 2016, The Second International Conference on Human and Social Analytics*. 2016. 29–32.
- Hoste, Veronique, Els Lefever, Stephan van der Waart van Gulik, et al. "TripleSent: a Triple Store of Events Associated with Their Prototypical Sentiment." *Proceedings of The Eighth International Conference on Information, Process, and Knowledge Management*. Venice, Italy: IARIA, 2016. 91–93.
- Hoste, Veronique, Walter Daelemans, Guy De Pauw, et al. "Natural Language Processing for Internet Security: The AMiCA Project." 2014.
- Hoste, Veronique. "The Mention-pair Model." *Anaphora Resolution : Algorithms, Resources and Applications*. Ed. Massimo Poesio, Roland Stuckardt, & Yannick Versley. Berlin Heidelberg: Springer-Verlag, 2016. 281–295.
- Jacobs, Gilles, Cynthia Van Hee, Chris Emmery, et al. "Cyberbullying Detection in Social Media." 2017.
- Jalalvand, Azarakhsh, Fabian Triefenbach, and Jean-Pierre Martens. "Continuous Digit Recognition in Noise: Reservoirs Can Do an Excellent Job!" *13th Annual Conference of the International Speech Communication Association, Proceedings*. Vol. 1–3. International Speech Communication Association (ISCA), 2012. 1802–1805.
- Jalalvand, Azarakhsh, Fabian Triefenbach, Kris Demuynck, et al. "Robust Continuous Digit Recognition Using Reservoir Computing." *COMPUTER SPEECH AND LANGUAGE* 30.1 (2015): 135–158.
- Jalalvand, Azarakhsh, Kris Demuynck, and Jean-Pierre Martens. "Feature Enhancement with a Reservoir-based Denoising Auto Encoder." *IEEE International Symposium on Signal Processing and Information Technology*. New York, NY, USA: IEEE, 2013. 227–232.
- Jalalvand, Azarakhsh, Kris Demuynck, and Jean-Pierre Martens. "Noise Robust Continuous Digit Recognition with Reservoir-based Acoustic Models." *International Symposium on Intelligent Signal Processing and Communication Systems ISPACS*. New York, NY, USA: IEEE, 2013. 204–209.
- Jalalvand, Azarakhsh, Kris Demuynck, Wesley De Neve, et al. "Design of Reservoir Computing Systems for Noise-robust Speech and Handwriting Recognition." *28th Conference on Graphics, Patterns and Images (Accepted in the Workshop of Theses and Dissertations (WTD)) Proceedings*. Sociedade Brasileira de Computação, 2015.
- Jalalvand, Azarakhsh, Wesley De Neve, Rik Van de Walle, et al. "Towards Using Reservoir Computing Networks for Noise-robust Image Recognition." *2016 INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN)* . Institute of Electrical and Electronics Engineers (IEEE), 2016. 1666–1672.
- Koleva, Mariya, Melissa Farasyn, Bart Desmet, et al. "An Automatic Part-of-speech Tagger for Middle Low German." *INTERNATIONAL JOURNAL OF CORPUS LINGUISTICS* 22.1 (2017): 108–141.
- Lavid, Julia, Jorge Arús, Bernard De Clerck, et al. "Creation of a High-quality, Register-diversified Parallel (English-Spanish) Corpus for Linguistic and Computational Investigations." *PROCEDIA* 198 (2015): n. pag.
- Lefever, Els, and Dennis J Folds, eds. "Proceedings of HUSO 2016, The Second International Conference on Human and Social Analytics." 2016 : n. pag.
- Lefever, Els, and Veronique Hoste. "A Classification-based Approach to Economic Event Detection in Dutch News Text." *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association (ELRA), 2016. 330–335.
- Lefever, Els, and Veronique Hoste. "Parallel Corpora Make Sense: Bypassing the Knowledge Acquisition Bottleneck for Word Sense Disambiguation." *INTERNATIONAL JOURNAL OF CORPUS LINGUISTICS* 19.3 (2014): 333–367.

- Lefever, Els, and Veronique Hoste. "SemEval-2013 Task 10: Cross-lingual Word Sense Disambiguation." Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2 : Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013). Association for Computational Linguistics (ACL), 2013. 158–166.
- Lefever, Els, Iris Hendrickx, and Antal van den Bosch. "Very Quaffable and Great Fun: Applying NLP to Wine Reviews." Computational Linguistics in the Netherlands. 2015.
- Lefever, Els, Liesbeth Allein, and Gilles Jacobs. "In the Eye of the Beer-holder : Lexical Descriptors of Aroma and Taste Sensations in Beer Reviews." HUSO 2017 : The Third International Conference on Human and Social Analytics. IARIA, 2017. 25–29.
- Lefever, Els, Marjan Van de Kauter, and Veronique Hoste. "Evaluation of Automatic Hypernym Extraction from Technical Corpora in English and Dutch." LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION. Ed. Nicoletta Calzolari et al. European Language Resources Association (ELRA), 2014. 490–497.
- Lefever, Els, Marjan Van de Kauter, and Veronique Hoste. "HypoTerm Detection of Hypernym Relations Between Domain-specific Terms in Dutch and English." Ed. Pamela Faber & Marie-Claude L'Homme. TERMINOLOGY 20.2 (2014): 250–278.
- Lefever, Els, Veronique Hoste, and Martine De Cock. "Discovering Missing Wikipedia Inter-language Links by Means of Cross-lingual Word Sense Disambiguation." Computational Linguistics in the Netherlands (CLIN). 2012.
- Lefever, Els, Veronique Hoste, and Martine De Cock. "Discovering Missing Wikipedia Inter-language Links by Means of Cross-lingual Word Sense Disambiguation." LREC 2012 : Eight International Conference on Language Resources and Evaluation. Ed. Nicoletta Calzolari et al. Paris, France: European Language Resources Association (ELRA), 2012. 841–846.
- Lefever, Els, Veronique Hoste, and Martine De Cock. "Five Languages Are Better Than One: An Attempt to Bypass the Data Acquisition Bottleneck for WSD." Lecture Notes in Computer Science. Ed. Alexander Gelbukh. Vol. 7816. Berlin, Germany: Springer, 2013. 343–354.
- Lefever, Els. "A Hybrid Approach to Domain-independent Taxonomy Learning." APPLIED ONTOLOGY 11.3 (2016): 255–278.
- Lefever, Els. "LT3: a Multi-modular Approach to Automatic Taxonomy Construction." Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). Association for Computational Linguistics, 2015. 944–948.
- Lefever, Els. "ParaSense: Parallel Corpora for Word Sense Disambiguation." 2012 : n. pag.
- Leijten, Mariëlle, Lieve Macken, Veronique Hoste, et al. "From Character to Word Level: Enabling the Linguistic Analyses of Inputlog Process Data." Proceedings of the EACL 2012 Workshop on Computational Linguistics and Writing. Association for Computational Linguistics (ACL), 2012.
- Leman, Marc, Dirk Moelants, Matthias Varewyck, et al. "Activating and Relaxing Music Entrain the Speed of Beat Synchronized Walking." Ed. Joel Snyder. PLOS ONE 8.7 (2013): n. pag.
- Macken, Lieve, and Arda Tezcan. "Dutch Compound Splitting for Bilingual Terminology Extraction." Multi-word Units in Machine Translation and Translation Technology. Ed. Ruslan Mitkov et al. John Benjamins, 2016.
- Macken, Lieve, Els Lefever, and Veronique Hoste. "TEXSIS: Bilingual Terminology Extraction from Parallel Corpora Using Chunk-based Alignment." TERMINOLOGY 19.1 (2013): 1–30.
- Macken, Lieve, Orphée De Clercq, Bart Desmet, et al. "Dutch Parallel Corpus En SoNaR." Beschouwingen Uit Een Talenhuis : Opstellen over Onderwijs En Onderzoek in De Vakgroep Vertalen, Tolken En Communicatie Aangeboden Aan Rita Godyns. Ed. Stefaan Evenepoel, Patrick Goethals, & Lieve Jooken. Gent: Academia Press, 2014. 163–170.
- Macken, Lieve, Veronique Hoste, Mariëlle Leijten, et al. "From Keystrokes to Annotated Process Data: Enriching the Output of Inputlog with Linguistic Information." LREC 2012 : Eight International Conference on Language Resources and Evaluation. Ed. Nicoletta Calzolari et al. Paris, France: European Language Resources Association (ELRA), 2012. 2224–2229.
- Macken, Lieve. "Hoe Een Computer Zelf Leert Vertalen." DIXIT, TIJDSCHRIFT OVER TAAL- EN SPRAAKTECHNOLOGIE (2017): 7–8.
- Middag, Catherine, Renee Clapham, Rob van Son, et al. "Robust Automatic Intelligibility Assessment Techniques Evaluated on Speakers Treated for Head and Neck Cancer." COMPUTER SPEECH AND LANGUAGE 28.2 (2014): 467–482.
- Moerman, Mieke, Jean-Pierre Martens, and Philippe jonckere. "Multidimensional Assessment of Strongly Irregular Voices Such as in Substitution Voicing and Spasmodic Dysphonia: A Compilation of Own Research." LOGOPEDICS PHONIATRICALS VOCOLOGY 40.1 (2015): 24–29.

- Montero Perez, Maribel, Hans Paulussen, Lieve Macken, et al. "From Input to Output: The Potential of Parallel Corpora for CALL." *LANGUAGE RESOURCES AND EVALUATION* 48.1 (2014): 165–189.
- Mourisse, Dieter, Els Lefever, Nele Verbiest, et al. "SBFC: An Efficient Feature Frequency-based Approach to Tackle Cross-lingual Word Sense Disambiguation." *Lecture Notes in Computer Science*. Ed. Petr Sojka et al. Vol. 7499. Berlin, Germany: Springer, 2012. 248–255.
- O'Donnell, Fionntán, Fabian Triefenbach, Jean-Pierre Martens, et al. "Effects of Architecture Choices on Sparse Coding in Speech Recognition." *Lecture Notes in Computer Science*. Vol. 7552. Berlin, Germany: Springer, 2012. 629–636.
- Oostdijk, Nelleke, Martin Reynaert, Veronique Hoste, et al. "SoNaR Nieuw Media Corpus." 2014 : n. pag.
- Oostdijk, Nelleke, Martin Reynaert, Veronique Hoste, et al. "SoNaR-1." 2014 : n. pag.
- Oostdijk, Nelleke, Veronique Hoste, Franciska de Jong, et al. "SoNaR-500." 2014 : n. pag.
- Paulussen, Hans , Lieve Macken, Willy Vandeweghe, et al. "Dutch Parallel Corpus: a Balanced Parallel Corpus for Dutch-English and Dutch-French." *Essential Speech and Language Technology for Dutch : Results by the STEVIN-programme*. Ed. Peter Spyns & Jan Odijk. Berlin, Germany: Springer, 2013. 185–199.
- Pauwels, Johan, and Jean-Pierre Martens. "Combining Musicological Knowledge About Chords and Keys in a Simultaneous Chord and Local Key Estimation System." Ed. Darell Conklin, Rafael Ramirez, & José M. Iñesta. *JOURNAL OF NEW MUSIC RESEARCH* 43.3 (2014): 318–330.
- Pelemans, Joris , Kris Demuynck, Hugo Van Hamme, et al. "Improving N-Gram Probability Estimates by Compound-head Clustering." *International Conference on Acoustics Speech and Signal Processing ICASSP*. NEW YORK: IEEE, 2015. 5221–5225.
- Pelemans, Joris, Kris Demuynck, and Patrick Wambacq. "A Layered Approach for Dutch Large Vocabulary Continuous Speech Recognition." *2012 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)*. IEEE, 2012. 4421–4424.
- Pelemans, Joris, Kris Demuynck, and Patrick Wambacq. "Dutch Automatic Speech Recognition on the Web: Towards a General Purpose System." *13TH ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION 2012 (INTERSPEECH 2012)*, VOLS 1-3. 2012.
- Pelemans, Joris, Kris Demuynck, Hugo Van Damme, et al. "The Effect of Word Similarity on N-gram Language Models in Northern and Southern Dutch." *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 4 (2014): 91–104.
- Pelemans, Joris, Kris Demuynck, Van hammeHugo, et al. "Coping with Language Data Sparsity: Semantic Head Mapping for Compound Words." *International Conference on Acoustics Speech and Signal Processing ICASSP*. IEEE, 2014. 141–145.
- Pelemans, Joris, Kris Demuynck, Van hammeHugo, et al. "Improving N-gram Probabilities by Compound-head Clustering." *International Conference on Acoustics, Speech, and Signal Processing*. 2015. 5221–5225.
- Pelemans, Joris, Kris Demuynck, Van hammeHugo, et al. "Speech Recognition Web Services for Dutch." *LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION*. 2014. 3041–3044.
- Pelemans, Joris, Tom Vanallemeersch, Kris Demuynck, et al. "Efficient Language Model Adaptation for Automatic Speech Recognition of Spoken Translations." *16th Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)*, Vols 1-5. Baixas, France: International Speech Communication Association (ISCA), 2015. 2262–2266.
- Pelemans, Joris, Tom Vanallemeersch, Lyan Verwimp, et al. "Language Model Adaptation for ASR of Spoken Translations Using Phrase-based Translation Models and Named Entity Models." *International Conference on Acoustics, Speech, and Signal Processing*. 2016. 5985–5989.
- Pontiki, Maria, Dimitris Galanis, Haris Papageorgiou, et al. "SemEval-2016 Task 5 : Aspect Based Sentiment Analysis." *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. Association for Computational Linguistics, 2016. 19–30.
- Réveil, Bert, Jean-Pierre Martens, and Henk van den Heuvel. "Improving Proper Name Recognition by Means of Automatically Learned Pronunciation Variants." *SPEECH COMMUNICATION* 54.3 (2012): 321–340.
- Réveil, Bert, Jean-Pierre Martens, Henk van de Heuvel, et al. "Lexical Modeling for Proper Name Recognition in Autonomata Too." *Essential Speech and Language Technology for Dutch*. Ed. Peter Spyns & Jan Odijk. Heidelberg, Germany: Springer, 2013. 251–27.
- Réveil, Bert, Kris Demuynck, and Jean-Pierre Martens. "An Improved Two-stage Mixed Language Model Approach for Handling Out-of-vocabulary Words in Large Vocabulary Continuous Speech Recognition." *COMPUTER SPEECH AND LANGUAGE* 28.1 (2014): 141–162.
- Reynaert, Martin, Ineke Schuurman, Veronique Hoste, et al. "Beyond SoNaR: Towards the Facilitation of Large Corpus Building Efforts." *LREC 2012 : Eight International Conference on Language Resources and Evaluation*.

- Ed. Nicoletta Calzolari et al. Paris, France: European Language Resources Association (ELRA), 2012. 2897–2904.
- Rigouts Terryn, Ayla, Lieve Macken, and Els Lefever. “Dutch Hypernym Detection: Does Decomposing Help?” Joint Second Workshop on Language and Ontology & Terminology and Knowledge Structures (LangOnto2 + TermiKS). European Language Resources Association (ELRA), 2016. 74–78.
- Rogova, Kseniya, Kris Demuynck, and Dirk Van Compernelle. “Automatic Syllabification Using Segmental Conditional Random Fields.” *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 3 (2013): 34–48.
- Schropp, Gwendolijn, Els Lefever, and Veronique Hoste. “A Combined Pattern-based and Distributional Approach for Automatic Hypernym Detection in Dutch.” *ATILA Workshop*. 2013.
- Schropp, Gwendolijn, Els Lefever, and Veronique Hoste. “A Combined Pattern-based and Distributional Approach for Automatic Hypernym Detection in Dutch.” *Proceedings of Recent Advances in Natural Language Processing*. Shoumen, Bulgaria: INCOMA, 2013. 593–600.
- Schulz, Sarah, Guy De Pauw, Orphée De Clercq, et al. “Multi-modular Text Normalization of Dutch User-generated Content.” *ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY* 7.4 (2016): n. pag.
- Shi, Yangyang, Martha Larson, Joris Pelemans, et al. “Integrating Meta-information into Recurrent Neural Network Language Models.” *SPEECH COMMUNICATION* 73 (2015): 64–80.
- Sinath, Tara N, Bhuvana Ramabhadran, David Nahamoo, et al. “Exemplar-based Processing for Speech Recognition.” *IEEE SIGNAL PROCESSING MAGAZINE* 29.6 (2012): 98–113.
- Smeuninx, Nils, Bernard De Clerck, Veronique Hoste, et al. “Sustainability Report Readability and Its Impact on Stakeholders’ Comprehension.” 2014.
- Smeuninx, Nils, Bernard De Clerck, Veronique Hoste, et al. “Sustainability Report Readability and Its Impact on Stakeholders’ Comprehension.” *Digital Trends in (Applied) Linguistics, Literature and Translation Studies*. Belgian Association of Anglicists in Higher Education, 2014.
- Smeuninx, Nils, Bernard De Clerck, Veronique Hoste, et al. “The Readability of Sustainability Reporting and Its Interaction with Company Performance.” *Association For Business Communication 2016 Conference*. 2016.
- Smeuninx, Nils, Orphée De Clercq, Veronique Hoste, et al. “Sustainability Report Readability Prediction: An NLP Expansion on the Tried-and-true Formulae.” *Computational Linguistics in the Netherlands* 26. 2015.
- Tezcan, Arda, Veronique Hoste, and Lieve Macken. “A Neural Network Architecture for Detecting Grammatical Errors in Statistical Machine Translation.” *THE PRAGUE BULLETIN OF MATHEMATICAL LINGUISTICS* 108 (2017): 133–145.
- Tezcan, Arda, Veronique Hoste, and Lieve Macken. “Detecting Grammatical Errors in Machine Translation Output Using Dependency Parsing and Treebank Querying.” *BALTIC JOURNAL OF MODERN COMPUTING* 4.2 (2016): 203–217.
- Tezcan, Arda, Veronique Hoste, and Lieve Macken. “SCATE Taxonomy and Corpus of Machine Translation Errors .” *Trends in E-tools and Resources for Translators and Interpreters*. Ed. Gloria Corpas Pastor & Isabel Durán-Muñoz. Vol. 45. Brill | Rodopi, 2017. 219–244.
- Tezcan, Arda, Veronique Hoste, and Lieve Macken. “UGENT-LT3 SCATE Submission for WMT16 Shared Task on Quality Estimation.” *Proceedings of the First Conference on Machine Translation, Volume 2: Shared Task Papers*. Association for Computational Linguistics, 2016. 843–850.
- Tezcan, Arda, Veronique Hoste, Bart Desmet, et al. “UGENT-LT3 SCATE System for Machine Translation Quality Estimation.” *Tenth Workshop on Statistical Machine Translation, Proceedings*. 2015.
- Triefenbach, Fabian, Azarakhsh Jalalvand, Kris Demuynck, et al. “Acoustic Modeling with Hierarchical Reservoirs.” *IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING* 21.11 (2013): 2439–2450.
- Triefenbach, Fabian, Azarakhsh Jalalvand, Kris Demuynck, et al. “Context-dependent Modeling and Speaker Normalization Applied to Reservoir-based Phone Recognition.” *14th Annual Conference of the International Speech Communication Association, Proceedings*. 2013. 3342–3346.
- Triefenbach, Fabian, Kris Demuynck, and Jean-Pierre Martens. “Improving Large Vocabulary Continuous Speech Recognition by Combining GMM-based and Reservoir-based Acoustic Modeling.” *IEEE Workshop on Spoken Language Technology, Proceedings*. Piscataway, NJ, USA: IEEE, 2012. 107–112.
- Triefenbach, Fabian, Kris Demuynck, and Jean-Pierre Martens. “Large Vocabulary Continuous Speech Recognition with Reservoir-based Acoustic Models.” *IEEE SIGNAL PROCESSING LETTERS* 21.3 (2014): 311–315.
- Van de Kauter, Marjan, Bart Desmet, and Veronique Hoste. “The Good, the Bad and the Implicit: a Comprehensive Approach to Annotating Explicit and Implicit Sentiment.” *LANGUAGE RESOURCES AND EVALUATION* 49.3 (2015): 685–720.
- Van de Kauter, Marjan, Bart Desmet, and Veronique Hoste. “The Good, the Bad and the Implicit: Annotating Polarity.” 2013.

- Van de Kauter, Marjan, Diane Breesch, and Veronique Hoste. "Fine-grained Analysis of Explicit and Implicit Sentiment in Financial News Articles." *EXPERT SYSTEMS WITH APPLICATIONS* 42.11 (2015): 4999–5010.
- Van de Kauter, Marjan, Geert Coorman, Els Lefever, et al. "LeTs Preprocess: The Multilingual LT3 Linguistic Preprocessing Toolkit." 2013.
- Van de Kauter, Marjan, Geert Coorman, Els Lefever, et al. "LeTs Preprocess: The Multilingual LT3 Linguistic Preprocessing Toolkit." *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 3 (2013): 103–120.
- Van de Velde, Stijn, Lieve Macken, Koen Vanneste, et al. "Technology for Large-scale Translation of Clinical Practice Guidelines: a Pilot Study of the Performance of a Hybrid Human and Computer-assisted Approach." *JMIR MEDICAL INFORMATICS* 3.4 (2015): n. pag.
- van den Heuvel, Henk, Jean-Pierre Martens, Gerrit Bloothoof, et al. "Resources Developed in the Autonomata Projects." *Essential Speech and Language Technology for Dutch*. Ed. Peter Spyns & Jan Odijk. Heidelberg, Germany: Springer, 2013. 61–78.
- Van den Steen, Leen, Gwen Van Nuffelen, Anke Aerts, et al. "Betrouwbaarheid Van De Automatische Analyse Van Spraakverstaanbaarheid Van Laryngectomiepatiënten Binnen ASISTO." *LOGOPEDIE* 29.4 (2016): 85–88.
- van der Lek-Ciudin, Iulianna, Ayla Rigouts Terryn, Geert Heyman, et al. "Translators' Methods of Acquiring Domain-specific Terminology. Information Retrieval in Terminology Using Lexical Knowledge Patterns." 21st Symposium on Language for Specific Purposes, LSP 2017. 2017.
- van der Vliet, Hennie, and Els Lefever. "Hoe Goed Werkt Termextractie Voor Het Nederlands?" *Terminologie in Het Nederlandse Taalgebied*. Ed. Els Ruijsendaal & Cornelia Wermuth. Gent: Academia Press, 2016. 47–74.
- Van Hee, Cynthia, Ben Verhoeven, Els Lefever, et al. "Data Set Construction and Exploratory Experiments for Cyberbullying Detection." 2014.
- Van Hee, Cynthia, Ben Verhoeven, Els Lefever, et al. "Guidelines for the Fine-grained Analysis of Cyberbullying, Version 1.0." *LT3 Technical Report Series 2015* : n. pag.
- Van Hee, Cynthia, Ben Verhoeven, Els Lefever, et al. "Proof-of-concept Experiments for the Fine-grained Classification of Cyberbullying Events." *Computational Linguistics in the Netherlands*. 2015.
- Van Hee, Cynthia, Celine Verleye, and Els Lefever. "Analysing Emotions in Social Media Coverage on Paris Terror Attacks : a Pilot Study." *Proceedings of HUSO 2016, The Second International Conference on Human and Social Analytics*. Ed. Els Lefever & Dennis J Folds. IARIA, 2016. 33–37.
- Van Hee, Cynthia, Els Lefever, and Veronique Hoste. "Exploring the Realization of Irony in Twitter Data." *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association (ELRA), 2016. 1795–1799.
- Van Hee, Cynthia, Els Lefever, and Veronique Hoste. "Guidelines for Annotating Irony in Social Media Text, Version 2.0." *LT3 Technical Report Series 2016* : n. pag.
- Van Hee, Cynthia, Els Lefever, and Veronique Hoste. "LT3: Sentiment Analysis of Figurative Tweets: Piece of Cake #NotReally." *SemEval : 9th International Workshop on Semantic Evaluations at NaacL 2015, Proceedings*. Denver, Colorado: Association for Computational Linguistics, 2015. 684–688.
- Van Hee, Cynthia, Els Lefever, and Veronique Hoste. "Monday Mornings Are My Fave : #not Exploring the Automatic Recognition of Irony in English Tweets." *Proceedings of COLING 2016, 26th International Conference on Computational Linguistics*. Ed. Nicoletta Calzolari, Yuji Matsumoto, & Rashmi Prasad. ACL, 2016. 2730–2739.
- Van Hee, Cynthia, Els Lefever, Ben Verhoeven, et al. "Automatic Detection and Prevention of Cyberbullying." *International Conference on Human and Social Analytics, Proceedings*. Ed. Pascal Lorenz & Christian Bourret. IARIA, 2015. 13–18.
- Van Hee, Cynthia, Els Lefever, Ben Verhoeven, et al. "Detection and Fine-grained Classification of Cyberbullying Events." *Proceedings of Recent Advances in Natural Language Processing, Proceedings*. Ed. Galia Angelova, Kalina Bontcheva, & Ruslan Mitkov. 2015. 672–680.
- Van Hee, Cynthia, Marjan Van de Kauter, Orphée De Clercq, et al. "LT3: Sentiment Classification in User-generated Content Using a Rich Feature Set." *Proceedings of the 8th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2014. 406–410.
- Vandevoorde, Lore, Els Lefever, Koen Plevoets, et al. "A Corpus-based Study of Semantic Differences in Translation : the Case of Inchoativity in Dutch." *TARGET-INTERNATIONAL JOURNAL OF TRANSLATION STUDIES* 29.3 (2017): 388–415.
- Vandevoorde, Lore, Koen Plevoets, Els Lefever, et al. "A Cognitive-translational Approach to Meaning: Visualizing Semantic Structure Through Translational Corpora." *Translation in Transition, Abstracts*. 2015.
- Vandevoorde, Lore, Pauline De Baets, Els Lefever, et al. "Distributional and Translational Solutions to the Visualization of Semantic Differences Between Translated and Non-translated Dutch." *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics*. Tübingen: Gerhard Jäger, 2016.

- Vanopstal, Klaar, Els Lefever, and Veronique Hoste. "Does Size Matter? A Comparison of a Large Web Corpus and a Smaller Focused Corpus for Medical Term Extraction." 2014 : 1612–1612.
- Vanopstal, Klaar, Lieve Macken, Els Lefever, et al. "Terminologie: Op Het Snijvlak Van Ambacht En Technologie." *Beschouwingen Uit Een Talenhuis : Opstellen over Onderwijs En Onderzoek in De Vakgroep Vertalen, Tolken En Communicatie Aangeboden Aan Rita Godyns*. Ed. S Evenepoel, P Goethals, & L Jooken. Gent: Academia Press, 2014. 179–189.
- Varewyck, Matthias, Jean-Pierre Martens, and Marc Leman. "Musical Meter Classification with Beat Synchronous Acoustic Features, DFT-based Metrical Features and Support Vector Machines." *JOURNAL OF NEW MUSIC RESEARCH* 42.3 (2013): 267–282.
- Verwimp, Lyan, Brecht Desplanques, Kris Demuynck, et al. "STON Efficient Subtitling in Dutch Using State-of-the-art Tools." *17TH ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (INTERSPEECH 2016), VOLS 1-5: UNDERSTANDING SPEECH PROCESSING IN HUMANS AND MACHINES*. Ed. Nelson Morgan et al. PO Box 100, Rundle Mall SA 5000, Australia: USB Production by Causal Productions Pty Ltd, 2016. 780–781.
- Wambacq, Patrick, Kris Demuynck, and Dirk Van Compernelle. "SPRAAK: Speech Processing, Recognition and Automatic Annotation Kit." *Essential Speech and Language Technology for Dutch*. Ed. Peter Spyns & Jan Odijk. Springer Berlin Heidelberg, 2013. 95–113.
- Wang, Yujun, Gemmeke Jort, Kris Demuynck, et al. "Missing Data Solutions for Robust Speech Recognition." *Essential Speech and Language Technology for Dutch*. Ed. Peter Spyns & Jan Odijk. Springer Berlin Heidelberg, 2013. 289–303.
- Zhang, Xueru, Kris Demuynck, and Van hammeHugo. "Latent Variable Speaker Adaptation of Gaussian Mixture Weights and Means." *2012 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)*. IEEE, 2012. 4349–4352.
- Zhang, Xueru, Kris Demuynck, and Van hammeHugo. "Rapid Speaker Adaptation in Latent Speaker Space with Non-negative Matrix Factorization." *SPEECH COMMUNICATION* 55.9 (2013): 893–908.
- Zhang, Xueru, Kris Demuynck, Dirk Van Compernelle, et al. "Subspace-GMM Acoustic Models for Under-resourced Languages: Feasibility Study." *Spoken Languages Technologies for Under-resourced Languages, Proceedings*. 2012. 1–4.

D. VRIJE UNIVERSITEIT BRUSSEL - OVERZICHT PUBLICATIES

- Beuls, K. & Loeckx, J., "Steps towards intelligent MOOCs: A case study for learning counterpoint", 2015 Music Learning With Massive Open Online Courses. Steels, L. (ed.). Amsterdam: IOS Press, p. 119-144 25 p. 9. (The Future of Learning; vol. 6)
- Beuls, K. & Steels, L., "Agent-Based Models of Strategies for the Emergence and Evolution of Grammatical Agreement", 2013 In : PLoS ONE. 8, 3, 15 p., e58960
- Beuls, K. & Steels, L., "Approaches to the verb phrase in Fluid Construction Grammar", 2017 (Accepted/In press) In : Constructions and Frames. 9, 2, p. 175-177
- Beuls, K. & Van Eecke, P., "Evaluating Computational Construction Grammars", 2017 CLIN-27: Book of Abstracts. p. 14-14 1 p.
- Beuls, K. & Van Eecke, P., "The Role of Morphological Marking in Reducing Semantic Ambiguity", 2017 The 30th Annual CUNY Conference on Human Sentence Processing. p. 561-562 2 p.
- Beuls, K. & Van Trijp, R., "Computational Construction Grammar and Constructional Change", 2016 In : Belgian Journal of Linguistics. 30, p. 1-13 13 p.
- Beuls, K. (ed.) "Luc Steels - Growing Seeds: Cutting-edge research & mentorship since 1982", 18 Sep 2017 Antwerp: Berger & Montag Publishers. 255 p.
- Beuls, K., "An Open-ended Construction-based Grammar for Spanish verb conjugation", 2017 (Accepted/In press) In : Constructions and Frames. 9, 2, p. 276-299
- Beuls, K., "A constructionist approach to student modeling: tracing a student's constructions through an agent-based tutoring architecture", Dec 2013 EUROCALL Conference proceedings: Learning from the Past, Looking to the Future.. p. 45-50
- Beuls, K., "An open-ended construction-based grammar for the Spanish verb phrase", 17 Dec 2015 Morphology Days 2015. Leuven: Instituut voor Mediastudies (KU Leuven), p. 1 1 p.
- Beuls, K., "Grammatical Error Diagnosis in Fluid Construction Grammar: A Case Study in L2 Spanish Verb Morphology", 2014 In : Computer Assisted Language Learning. 27, 3, p. 246-260
- Beuls, K., "Handling Scope in Fluid Construction Grammar: A Case Study for Spanish Modals", 2012 Computational Issues in Fluid Construction Grammar. Steels, L. (ed.). Springer, p. 123-142 20 p. (Lecture Notes in Artificial Intelligence; vol. 7249)
- Beuls, K., "Inflectional patterns as constructions: Spanish verb morphology in Fluid Construction Grammar", 2012 In : Constructions and Frames. 4, 2, p. 231-250 20 p.
- Beuls, K., Ferdinand, V. & Kanwai, J., "Self-Organisation of a Learnable Language: What Happens When an Alien Language Replicates through Human Brains", 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL), 2012.
- Beuls, K., Knight, Y. & Spranger, M., "Russian verbs of motion and their aspectual partners in Fluid Construction Grammar", 2017 (Accepted/In press) In : Constructions and Frames. 9, 2, p. 300-318
- Beuls, K., Steels, L. & Höfer, S., "The Emergence of Internal Agreement Systems", 2012 Experiments in Cultural Language Evolution. Steels, L. (ed.). John Benjamins Publishing Company, p. 233-256 24 p. (Advances in Interaction Studies; vol. 3)
- Beuls, K., Steels, L. & Van Eecke, P., "Strategies for Interactive Task Learning and Teaching", 2017 (Accepted/In press) Interactive Task Learning: Agents, Robots, and Humans Acquiring New Tasks through Natural Interaction. Gluck, K. & Laird, J. E. (eds.). MIT Press, (Ernst Strüngmann Forum)
- Beuls, K., Van Eecke, P. & Steels, L., "A Library of Meta-Level Diagnostics and Repairs for Fluid Construction Grammar", 2016 BNAIC 2016: Proceedings of the Twenty-Eight Benelux Conference on Artificial Intelligence. p. 224-226 3 p.
- Gerasymova, K., Spranger, M. & Beuls, K., "A Language Strategy for Aspect: Encoding Aktionsarten through Morphology", 2012 Experiments in Cultural Language Evolution. Steels, L. (ed.). John Benjamins Publishing Company, p. 257-276 20 p. (Advances in Interaction Studies; vol. 3)
- Hoorens, S., Beuls, K. & Van Eecke, P., "Constructions at Work! Visualising Linguistic Pathways for Computational Construction Grammar", 8 Nov 2017 Proceedings of the 29th Benelux Conference on Artificial Intelligence. p. 224-237
- Pijpops, D. & Beuls, K., "Agent-gebaseerde modellering in de historische taalkunde. Een model van regularisatiedruk op de Nederlandse werkwoorden", Dec 2015 In : Handelingen van de Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis. 69, p. 5-23 19 p.
- Pijpops, D., Beuls, K. & Van de Velde, F., "Fighting in broken formation. The competition between the Germanic strong ablaut classes and weak suffix inflection in an agent-based model." 3 Sep 2015 Shifting classes:

- Germanic strong and weak preterites and participles. Beuls, K., Bernard, D. C., Pijpops, D. & Van de Velde, F. (eds.). p. 440-441 2 p.
- Pijpops, D., Beuls, K. & Van de Velde, F., "The rise of the verbal weak inflection in Germanic: An agent-based model", 13 Nov 2015 In : Computational Linguistics in the Netherlands. 5, p. 81-102 23 p.
- Radulescu, R. & Beuls, K., "Modelling pronominal gender agreement in Dutch: From a syntactic to a semantic strategy", Dec 2016 In : Belgian Journal of Linguistics. 30, p. 219-250 10
- Radulescu, R-T. & Beuls, K., "A Multi-Agent Model Approach to Resemanticization in Pronominal Agreement in Dutch", 18 Dec 2015 Proceedings of the 26th Meeting of Computational Linguistics in the Netherlands (CLIN 26). Amsterdam: VU Amsterdam, p. 10-10 1 p.
- Santos Marques, T. F. & Beuls, K., "A Construction Grammar Approach for Pronominal Clitics in European Portuguese", 2016 Computational Processing of the Portuguese Language: 12th International Conference, PROPOR 2016, Tomar, Portugal, July 13-15, 2016, Proceedings. Springer Verlag, p. 239-244 5 p. (Lecture Notes in Artificial Intelligence)
- Santos Marques, T. F. & Beuls, K., "Evaluation Strategies for Computational Construction Grammars", Dec 2016 Proceedings of COLING 2016: the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan: The COLING 2016 Organizing Committee, p. 1137-1146 9 p.
- Spranger, M. & Beuls, K., "Referential Uncertainty and Word Learning in High-Dimensional, Continuous Meaning Spaces", 2016 6th International Conference on Development and Learning and on Epigenetic Robotics. p. 95-100 6 p.
- Steels, L. & Beuls, K., "How to explain the origins of complexity in language: A case study for agreement systems", 2017 Complexity in Language: Developmental and Evolutionary Perspectives. Mufwene, S., Coupé, C. & Pellegrino, F. (eds.). Cambridge: Cambridge University Press, p. 30-47 18 p. 2. (Cambridge Approaches to Language Contact)
- Van Eecke, P. & Beuls, K., "Meta-Layer Problem Solving for Computational Construction Grammar", 2017 The 2017 AAAI Spring Symposium Series. Palo Alto, California: AAAI Press, p. 258-265 8 p.
- Van Trijp, R., Steels, L., Beuls, K. & Wellens, P., "Fluid Construction Grammar: The New Kid on the Block", 2012 Proceedings of the 13th Conference of the European Chapter of the Association for computational Linguistics.
- Wellens, P., Beuls, K. & Van Trijp, R., "Diagnostics and Repairs in Fluid Construction Grammar", 2012 Language Grounding in Robots. Steels, L. & Hild, M. (eds.). Springer, p. 215-234
- Wellens, P., Van Trijp, R., Beuls, K. & Steels, L., "Fluid Construction Grammar for Historical and Evolutionary Linguistics", 2013 Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. p. 127-132 6 p.