

Een Kennisbank over Begrijpelijke Taal

1. Inleiding

Begrijpelijke taal – daar ziet iedereen de zin van in. Maar weten we ook wat een tekst begrijpelijk maakt, of onbegrijpelijk? 50 jaar na de publicatie van de eerste Nederlandse leesbaarheidformule (Douma, 1960) is er reden om de balans op te maken van het onderzoek naar begrijpelijkheid.

Dat onderzoek concentreerde zich aanvankelijk op de ontwikkeling van leesbaarheidsformules voor onderwijsteksten. Vanaf de zeventiger jaren van de vorige eeuw kwam er steeds meer kritiek op die benadering (Kintsch & Vipond, 1979; Anderson & Davison, 1988). Sinds die tijd werken taalkundigen, psychologen en onderwijskundigen op verschillende manieren aan begrijpelijkheid. Naast het theoretisch geïnspireerde begrijpelijkheidsonderzoek, en vaak enigszins los daarvan, is een stroom praktisch georiënteerd onderzoek op gang gekomen, bijvoorbeeld ten behoeve van de optimalisering van allerlei tekstgenres in de gezondheidscommunicatie.

Het resultaat van deze ontwikkelingen is dat de literatuur over begrijpelijkheid veelvoudig en onoverzichtelijk is. Het is lastig om toonaangevende tijdschriften te noemen. Een veelvoud aan tekstkenmerken wordt onderzocht en eenzelfde tekstkenmerk wordt vaak op uiteenlopende manieren gedefinieerd. Allerlei begripsmetingen worden naast elkaar gebruikt. Sommige studies onderzoeken minimale manipulaties op stimuli van minimale omvang, terwijl in andere studies volledige teksten grondig herschreven en (soms wat minder grondig) getest worden. Daar komt bij dat het onderzoek in verschillende talen, maar grotendeels in het Engels gepubliceerd wordt.

Dat de situatie onoverzichtelijk is, ervaren studenten ook. Ook na het volgen van cursussen op het gebied van tekstontwerp en begrijpelijkheid hebben zij moeite om zelfstandig

Samenvatting

Vaak moeten adviezen voor begrijpelijke taal het stellen zonder empirische ondersteuning. Dit artikel doet verslag van de ontwikkeling van de Kennisbank Begrijpelijke Taal, een database met 469 artikelen over empirisch begrijpelijkheidsonderzoek die zowel in het Nederlands als in het Engels beschikbaar is. De gebruiker kan deze artikelen bijvoorbeeld vinden door te zoeken op kenmerken als het onderzochte tekstgenre, de onderzochte boodschapkenmerken, en de gehanteerde begripsmaten. Hieronder is beschreven welk type publicaties is geselecteerd voor de kennisbank, hoe deze publicaties zijn gezocht, hoe ze zijn beschreven, en hoe het ontwerp van de kennisbank zelf is. Vervolgens geven we globaal overzicht van de inhoud van de kennisbank, en een voorbeeld van een zoektocht in de kennisbank aan de hand van de vraag: welk effect heeft de lijdende vorm op begrip? We sluiten af met enkele toekomstperspectieven.

relevante studies te vinden, ondanks de aanwezigheid van krachtige zoeksystemen. Evenzo hebben professionals op het gebied van tekst en communicatie in de dagelijkse praktijk moeite om kennis te verzamelen over het verband tussen tekstkenmerken en begrijpelijkheid. Zij hebben immers na hun opleiding het onderzoek meestal niet bijgehouden, zoals blijkt uit het onderzoek van Lentz e.a. (2000) onder professionele tekstschrijvers, en zullen daarom vaak de trefwoorden niet kennen waarmee de literatuur ontsloten kan worden.

Toen het NWO-programma Begrijpelijke Taal¹ van start ging, werd het daarom nuttig geacht om een verkenning uit te voeren van het beschikbare onderzoek. Deze verkenning heeft geresulteerd in een kennisbank voor onderzoekers, studenten en professionals op het gebied van taal en tekst: www.kennisbank-begrijpelijgetaal.nl. Aan dit project hebben wij van september 2010 tot maart 2011 gewerkt met een projectgroep bestaande uit onderzoekers van de Universiteit Utrecht en de Vrije Universiteit Amsterdam.² In dit artikel beschrijven we de achtergronden van deze kennisbank. Eerst beschrijven we de werkwijze bij het verzamelen van publicaties. Daarna gaan we in op de vraag hoe die publicaties beschreven zijn. Ten slotte gaan we wat dieper in op de inhoud van de kennisbank.

2. Welke publicaties zijn opgenomen in de kennisbank?

Een belangrijk deel van deze verkenning bestond uit het verzamelen van literatuur. Vier projectmedewerkers hebben in verschillende zoekmachines empirisch onderzoek naar Begrijpelijke Taal gezocht. De werkzaamheden werden aan het begin verdeeld in drie projecten. In project 1 ging het om psycholinguïstisch onderzoek naar begrijpelijkheid, in project 2 om verbeteringsgericht onderzoek uit de Document Design traditie, en in project 3 om de relatie tussen tekst, beeld en begrijpelijkheid.

De hoofdvraag in het psycholinguïstisch deelproject was de volgende: *Welke publicaties rapporteren over onderzoek naar de invloed van kenmerken van woorden, zinnen en teksten op de begrijpelijkheid van teksten?* Er is in dit deelproject gezocht naar experimenteel of correlatieel onderzoek naar het effect op begrijpelijkheid van:

- Woordmoeilijkheid;
- Zinsbouw;
- Tekstopbouw (kopjes, volgorde);
- Tekstvormgeving (zoals opsommingen, regellengte).

In de praktijk is het overgrote deel van de gevonden publicaties experimenteel van aard. Enkele voorbeelden van correlatieve studies zijn Meyer e.a. (1993) en Murray e.a. (1993). Omdat deze studies veruit in de minderheid zijn, is besloten om in de kennisbank een experimentele terminologie te hanteren. Zo wordt gesproken over ‘manipulaties’ wanneer het gaat om tekstvariabelen, terwijl daarvan in de schaarse correlatieve studies natuurlijk geen sprake is.

In het Document Design deelproject ging het om de vraag: *Welke publicaties rapporteren over onderzoek waarin reël bestaande teksten aantoonbaar begrijpelijker zijn gemaakt door veranderingen in woordmoeilijkheid, zinsbouw, tekstopbouw of tekstvormgeving?* Strikt genomen vallen deze studies in het ruime gebied dat ook is aangeduid in de psycholinguïstische vraag hierboven. Ze zijn in een apart project ondergebracht omdat we vermoedden dat studies met bestaande, authentieke teksten via andere zoekstrategieën gezocht zouden moeten worden dan

psycholinguïstische studies. Een aantal van deze studies is niet experimenteel in strikte zin. Regelmatig is de manipulatie meervoudig van aard, in die zin dat er meerdere boodschapkenmerken tegelijk gevarieerd worden.

In het project rond de relaties tussen tekst en beeld luidde de hoofdvraag als volgt: *Welke publicaties rapporteren over onderzoek naar factoren die invloed hebben op de begrijpelijkheid van tabellen, diagrammen, grafieken en illustraties?*

Kortom, de kennisbank bevat onderzoek waarin het effect van boodschapvariabelen op begrip empirisch onderzocht wordt. Er zijn honderden studies die aan dat criterium voldoen, maar er is ook erg veel onderzoek dat op het eerste gezicht misschien relevant lijkt en toch is uitgesloten. Hoe hebben we die selectie uitgevoerd? Hieronder bespreken we een aantal inhoudelijk-thematische criteria. Criteria wat betreft de kwaliteit van het onderzoek zijn niet gehanteerd. Dit omdat we een breed overzicht willen bieden. Overigens kunnen uit de beschrijving van de publicaties wel enige conclusies worden getrokken over de kwaliteit van het onderzoek.

Er is veel psychologisch en onderwijskundig onderzoek naar taalbegrip en leesvaardigheid waarin *geen sprake is van gemanipuleerde tekstkenmerken*. Ook in taalbeheersingsonderzoek komt dat voor, zoals in onderzoek naar de vraag in hoeverre lezers een bepaald genre begrijpen. Pander Maat e.a. (2009) rapporteren bijvoorbeeld over de begrijpelijkheid van hypothekinformatie, zonder tekstkenmerken te manipuleren. Daarnaast zijn strikt probleemopsporende studies uitgesloten en verslagen van ontwikkeltrajecten zonder tests. Ook studies die leesbaarheidsformules toepassen op teksten, die rapporteren welke informatie mensen zouden willen lezen of die alleen oordelen over begrijpelijkheid verzamelen, doen in onze definitie geen empirisch begripsonderzoek en blijven dus buiten de kennisbank.

Onderzoek naar het *effect van proefpersoonkenmerken* als leeftijd en voorkennis op leesbegrip is alleen opgenomen wanneer deze kenmerken samen met tekstkenmerken onderzocht worden. Ook onderzoek naar de *effecten van leesstrategieën* (zoals Salmeron e.a., 2010) of van de leessnelheid is niet in de kennisbank te vinden.

Verder zijn niet alle boodschapmanipulaties geaccepteerd. Onderzoek met *substantiële veranderingen van de inhoud in het tekstmateriaal* is niet opgenomen in de kennisbank. Het effect van korte toelichtingen is bijvoorbeeld nog meegenomen als tekstkenmerk, maar het invoegen van compleet nieuwe informatie niet. Daarom is bijvoorbeeld een studie naar verbeteringen van jury-instructies in de VS uitgesloten (Diamond & Levi, 1996). Wanneer twee tekstversies worden onderzocht en de ene is veel langer of korter dan de andere, doet de studie niet mee in de kennisbank (vergelijk Fuchs & Hippius, 2007). Maar ook psycholinguïstische studies vallen soms af door inhoudelijke manipulaties. Zo onderzoeken Moxey e.a. (2004, experiment 2 en 4) het lezen van verschillende vervolgen op zinnen met complexe subjecten als *Jan en Theo*. Het ene vervolg begint met het meervoudige *zij*, het andere met *hij*. Dat verschil is inhoudelijk: je kunt niet zeggen dat een globaal gelijke inhoud hier anders geformuleerd wordt. Daarom is deze studie is niet opgenomen.

Wat de *doelgroep* betreft onderzoeken we alleen teksten bestemd voor *leken* op het themadomein van de tekst, waartoe we ook studenten van een bepaald vak rekenen. De proefpersonen in het onderzoek moeten verder de taal van de tekst als *moedertaal* spreken. Dat is geen principiële maar een praktische keuze om haalbaarheidsredenen; we sluiten dus publicaties uit waarin gerapporteerd wordt over onderzoek naar problemen van Turkse immigranten in Duitsland met teksten in het Duits. Verder moeten de lezers in het onderzoek *12 jaar of ouder*

zijn. Onderzoek naar jongere lezers richt zich vaak mede op de verwerving van leesvaardigheid op de basisschool, en dat onderwerp laten we buiten beschouwing. De proefpersonen mogen verder geen specifieke *taal- of leesproblemen* hebben. Dat sluit onderzoek naar tekstbegrip van afasiepatiënten uit, maar onderzoek naar de invloed van tekstkenmerken op het leesgedrag van slechte lezers of laaggeletterden valt weer wel binnen de criteria.

Ook voor het *type* publicatie formuleerden we enkele criteria. De studie moet gepubliceerd zijn in een tijdschrift, congresbundel of boek. Scripties worden bijvoorbeeld niet opgenomen. Die zijn niet bij elke universiteit publiek toegankelijk. De allerbeste scripties komen vaak via een publicatie in een tijdschrift beschikbaar. Dat geldt ook voor nog ongepubliceerde rapporten die op het Internet circuleren.

Verder moet de publicatie verschenen zijn in *1980 of later*. De exacte plaatsing van zo'n tijdsdrempel is wat arbitrair, maar onze redenering was als volgt. Pas vanaf 1980 zijn zowel de tekstverwerkingspsychologie als het document design onderzoek echt tot wasdom gekomen. Tot die tijd wordt er aanzienlijk minder begrijpelijkheidsonderzoek gedaan dat relevant is voor tekstontwerp. Het merendeel van het klassieke leesbaarheidsonderzoek is minder relevant voor tekstontwerp, omdat de nadruk daar ligt op predictieve validiteit en niet op begripsprocessen en oorzaken van begripsproblemen.

We beperkten ons tot publicaties in het *Engels, Nederlands, Frans, Spaans of Duits*. In de praktijk zijn niet-Engelse treffers zeldzaam. Dat komt wellicht ook doordat vooral met Engelse trefwoorden is gezocht. Engelse trefwoorden leveren namelijk ook niet-Engelse publicaties op, maar het omgekeerde geldt niet. Zo levert het trefwoord *readability* in Scopus 4648 publicaties op, waarvan 163 in het Duits en 87 in het Frans. De term *Verständlichkeit* daarentegen levert 58 Duitse publicaties op, en *Lesbarkeit* 23. Daarnaast zijn er simpelweg veel meer publicaties in het Engels dan in andere talen omdat het Engels de dominante wetenschapstaal is en omdat in dit taalgebied veel relevant onderzoek is uitgevoerd.

Ten slotte werden *reviewartikelen* niet opgenomen in de kennisbank. Wel werden ze gebruikt als vindplaats voor mogelijke treffers.

Elke studie is slechts *eenmaal* in de kennisbank opgenomen. Wanneer experimenten twee maal worden gerapporteerd, wordt dus een van de publicaties buiten beschouwing gelaten. Zo wordt bijvoorbeeld Murray (1997) wel, maar Murray (1995) niet opgenomen. Wij kozen de meest recente publicatie, omdat daarin doorgaans een verband wordt gelegd met later uitgevoerde vervolgstudies. Anderzijds komen sommige artikelen als meerdere treffers in de kennisbank, omdat zij verschillende studies rapporteren. Voor elk van die studies is apart besloten of opname in de kennisbank wenselijk is.

Hoe zorgvuldig selectiecriteria ook geformuleerd worden, er blijven natuurlijk altijd twijfelgevallen bestaan. We noemen er enkele. In het onderzoek van McNamara en haar collega's naar cohesie van leerteksten is regelmatig sprake van ingrijpende manipulaties. Zo verlenen O'Reilly en McNamara (2007) een tekst van 650 naar 901 woorden. Ze doen dat niet alleen door het vervangen van voornaamwoorden door naamwoorden, het invoegen van connectieven en signaalzinnen en het invoegen van woorden die conceptuele overlap met de voorgaande zin opleveren, maar ook door het toevoegen van themazinnen en 'descriptive elaborations'. Met name de laatste categorie lijkt inhoudelijke ingrepen te kunnen bevatten; de overige ingrepen kunnen gezien worden als het expliciteren van structuur. Een ander twijfelgeval dat we toch hebben opgenomen is het onderzoek naar 'verleidelijke details' door

bijvoorbeeld Schraw (1998). In dit onderzoek worden inhoudelijke toevoegingen gedaan, maar die betreffen niet de hoofdlijn van het verhaal. Ten slotte is er de onderzoekslijn van Sadoski en zijn collega's naar tekstconcreetheid (bijvoorbeeld Sadoski e.a., 2000). Hoewel de concreetheid bij Sadoski eerder de tekstinhoud dan de formulering betreft, hebben we zijn werk om twee redenen opgenomen. Ten eerste wordt gepoogd het effect van concreetheid te onderscheiden van de effecten van meer traditionele leesbaarheidskenmerken. Ten tweede wordt het werk uitdrukkelijk geplaatst in de discussie over document design en zouden veel gebruikers van de kennisbank het missen wanneer het zou ontbreken. We geven toe dat dit laatste argument niet erg principieel van aard is.

3. Hoe zijn publicaties gezocht?

Het zoeken naar wetenschappelijke publicaties verliep in drie fasen. We hebben allereerst een zoekstrategie bepaald, vervolgens combinaties van zoektermen gekozen en ten slotte de relevante publicaties geselecteerd.

3.1 Fase 1: het bepalen van de zoekstrategie Voor het onderzoek naar tekst en begrip zijn de volgende zoekmachines geprobeerd:

- CSA Linguistics & Language Behavior Abstracts (LLBA)
- MLA International Bibliography
- Web of Science
- Scopus
- Google Scholar
- PsycINFO

In deze zoekmachines hebben we in het begin meer algemene zoektermen ingevoerd. Bijvoorbeeld 'text AND comprehension'. In alle zoekmachines leverde dit veel resultaten op, maar weinig treffers. Met 'treffer' bedoelen we: een studie die voldoet aan de eisen. Naarmate we de zoektermen verder inperkten door woorden toe te voegen, leverde het zoeken meer relevante treffers op.

Om de kwaliteit van de zoekmachine te testen, hadden we een aantal sleutelpublicaties bepaald, die in elk geval in de output als treffer tevoorschijn zouden moeten komen. Dit waren publicaties waar doorgaans naar verwezen wordt in artikelen over bijvoorbeeld het domein van tekst en begrip. Uiteindelijk hebben we naar aanleiding van de resultaten in de eerste zoektochten gekozen voor een drietal zoekmachines:

- *Scopus* biedt goede resultaten, zoekt in veel verschillende bronnen, en is gemakkelijk in gebruik.
- *Google Scholar* biedt erg uitgebreide resultaten, met de meest relevante bovenaan; maar noodzaakt wel tot verfijnder gebruik van zoektermen.
- *PsycINFO* biedt vooral andere resultaten dan de twee andere zoekmachines.

De volgende zoekmachines zijn hiernaast nog gebruikt voor het onderzoek naar tekst en beeld:

- CSA Linguistics & Language Behavior Abstracts (LLBA)
- Web of Science
- Bibliografie van de Nederlandse Taal- en Literatuurwetenschap (BNTL)

Ondanks de scheiding van de zoekwerkzaamheden in drie aparte projecten (psycholinguïstiek, document design en visueel), was er vanzelfsprekend overlap binnen de gevonden resultaten. Om geen dubbele treffer in de kennisbank te krijgen, hebben we onderling afspraken gemaakt over wie welke treffer zou behandelen in twijfelgevallen.

3.2 Fase 2: Beschrijving van de zoekacties In de eerste fase bleek dat zoekmachines vaak veel resultaten leveren die bij nader inzien onbruikbaar zijn. Dat is natuurlijk ook een reden voor de ontwikkeling van deze kennisbank. Onderzoekers hebben immers baat bij het vinden van publicaties op het gebied van begrijpelijke taal, maar zien door de vele resultaten in zoekmachines wellicht door de bomen het bos niet meer. Om het zoeken te vergemakkelijken hebben we een schema opgesteld met trefwoorden en mogelijke trefwoordcombinaties. In het schema worden de trefwoorden in drie klassen verdeeld:

- Termen die verwijzen naar het onderzochte effect
- Termen die verwijzen naar onderzochte tekstenkenmerken
- Termen die verwijzen naar genre-aanduidingen (bv. *patient leaflets*) en/of tekstcomponenten (woord, zin, enzovoort).

Zo is er een zoekactie uitgevoerd met de trefwoorden *language comprehension* (effectterm) *AND word familiarity* (tekstkenmerk) in zoekmachine Scopus. Van de 39 resultaten gingen er drie over het begrijpen van spreekwoorden. Deze zijn opgenomen in de kennisbank.

Om de zoekresultaten bewerkbaar te houden, hanteerden we een strategie waarbij we nieuwe zoektermen toevoegden als er meer dan 200 resultaten waren. Wie bijvoorbeeld met de termen *text AND comprehension* zoekt in Scopus krijgt 6.651 resultaten. Maar met toevoeging van *experiment* is het aantal resultaten 983. En met een extra toevoeging van het genre als trefwoord, bijvoorbeeld 'news', zijn er nog maar 11 resultaten. Ook voor de zoekmachine Google Scholar was die strategie noodzakelijk, vanwege de grote hoeveelheid resultaten. Dit leidde tot extreem lange zoekstrings. Op die manier kwamen vaak belangrijke treffer naar voren, dus deze strategie bleek goed te werken.

Op de website is een document te vinden met een nadere toelichting op de werkwijze. Er werd altijd een combinatie gebruikt van een kenmerk en het effect daarvan op de begrijpelijkheid. Hetzelfde document geeft de complete lijst met alle 340 zoektermen van het psycholinguïstisch onderzoek en alle 262 zoektermen van het op Document Design gerichte onderzoek, inclusief de zoekmachines die gebruikt zijn bij de termen. Bovendien is er een beschrijving te vinden van de zoekacties naar publicaties over de relatie tussen tekst en beeld.

Voorbeelden van zoektermen zijn:

- word difficulty AND text comprehension,
- signals AND reading comprehension,
- effect of word repetition, comprehension AND news reports,
- understanding AND legal text AND improve,
- comprehension history text AND experiment.

Naast deze systematische manier van termen invoeren in zoekmachines, zijn ook enkele specifieke tijdschriften en congresbundels handmatig of digitaal doorzocht (*Document Design* cq. *Information Design Journal* en *Studies in Taalbeheersing*). Verder is per tekstkenmerk en per genre telkens vooruit en achteruit gekeken. We keken vooruit door bij elk oudste artikel in

een verzameling treffers na te gaan welke artikelen later een referentie opnamen naar dit oudste artikel. We keken achteruit door bij elk nieuwste artikel van een groep de referenties naar eerdere literatuur te controleren. Soms leidde dat tot treffers die we nog niet met de zoekmachines gevonden hadden. Tot slot zijn enkele sleutelartikelen geselecteerd, voor deze artikelen hebben we dezelfde werkwijze uitgevoerd met vooruit en achteruit kijken.

3.3 Fase 3: Selectie van artikelen Hoeveel artikelen uit één zoekstring komen nu eigenlijk in de Kennisbank Begrijpelijke Taal terecht? Een voorbeeld van een zoekstring uit het psycholinguïstische project is: *Word frequency AND text comprehension*. In zoekmachine Scopus komen met deze termen op 1 maart 2011 85 resultaten naar voren van na 1979, waarvan er uiteindelijk 3 in de Kennisbank terecht zijn gekomen:

- Lachman, R. (1989). Comprehension aids for on-line reading of expository text. *Human Factors*, 31(1), 1-15.
- Lenzner, T., Kaczmirek, L., & Lenzner, A. (2010). Cognitive burdens of survey questions and response times: A psycholinguistic experiment *Applied Cognitive Psychology*, 24(7), 1003-1020.
- Williams, R. S., & Morris, R. K. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, 16(1-2), 312-339.

Welke treffers uit deze 85 zijn er zoal niet geschikt voor de kennisbank? Een paar voorbeelden:

- Lawrie, D., Binkley, D., & Morrell, C. (2010). Normalizing source code vocabulary. *Proceedings - Working Conference on Reverse Engineering, WCRE, 2010*, 3-12.
Dit artikel gaat over het consistent maken van het vocabulair in de verschillende documenten rond een software-ontwikkelingsproject.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, 27(1), 57-86.
Dit artikel gaat over automatische analyse van de schrijfkwaliteit zoals die uit een tekst blijkt.
- Sabbahi, D. A., Lawrence, H. P., Limeback, H., & Rootman, I. (2009). Development and evaluation of an oral health literacy instrument for adults. *Community Dentistry and Oral Epidemiology*, 37(5), 451-462.
Dit artikel gaat over het ontwerpen van een nieuw instrument om de gezondheidsgeletterdheid te meten.
- Schad, D. J., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly shuffled text. *Vision Research*, 50(23), 2600-2616.

Dit artikel gaat over hoe de ogen bewegen door een tekst waarvan alle woorden door elkaar zijn gegooid, en hoe dat verschilt van oogbewegingen door gewone teksten. Dit artikel komt niet in aanmerking omdat deze 'scrambling'-manipulatie geen alternatieve tekst produceert, maar een verzameling woorden.

Een voorbeeld van een zoekstring uit het document designproject is: *Readability AND "Informed consent information"*. Op deze termcombinatie komen in Scopus 85 resultaten van na 1979, waarvan er 5 uiteindelijk in de Kennisbank Begrijpelijke Taal terechtkomen:

- Bjørn, E., Rossel, P., & Holm, S. (1999). Can the written information to research subjects be improved? An empirical study. *Journal of Medical Ethics*, 25(3), 263-267.

- Kang e.a. (2009). Informed consent recall and comprehension in orthodontics: Traditional vs. improved readability and processability methods. *American Journal of Orthodontics and Dentofacial Orthopedics*, 136(4), 488.e1-488.e13.
- Knapp, P., Raynor, D. K., Silcock, J., & Parkinson, B. (2009). Performance-based readability testing of participant information for a phase 3 IVF trial. *Trials* 10, art. no. 1745, 79.
- Paris, A., Brandt, C., Cornu, C., Maison, P., Thalamas, C., & Cracowski, J. (2010). Informed consent document improvement does not increase patients' comprehension in biomedical research. *British Journal of Clinical Pharmacology*, 69(3), 231-237.
- Tait, A. R., Voepel-Lewis, T., Malviya, S., & Philipson, S. J. (2005). Improving the readability and processability of a pediatric informed consent document: Effects on parents' understanding. *Archives of Pediatrics Adolescent Medicine*, 159(4), 347-352.

Welke publicaties komen niet in de kennisbank? Wederom enkele voorbeelden.

- Barrio-Cantalejo, I. M., Simón-Lorda, P., Sánchez Rodríguez, C., Molina-Ruiz, A., Tamayo-Velázquez, M. I., Suess, A., et al. (2009). Cross-cultural adaptation and validation of the picker patient experience questionnaire-15 for use in the spanish population. [Adaptación transcultural y validación del Picker Patient Experience Questionnaire-15 para su uso en población española] *Revista De Calidad Asistencial*, 24(5), 192-206.

Dit artikel gaat over de aanpassing van een vragenlijst over patiënttevredenheid.

- Kellen, E., Van Steen, A., Bleyen, L., Coelst, M. E., Goossens, M., Martens, P., et al. (2010). Carefully weighing the advantages and disadvantages of the screening program for breast cancer in Flanders. [De voor- en nadelen van het systematisch bevolkingsonderzoek naar borstkanker in Vlaanderen zorgvuldig afgewogen] *Tijdschrift Voor Geneeskunde*, 66(18), 845-851.

De titel van dit artikel maakt duidelijk dat het niets met taal te maken heeft.

- Ménoni, V., Noël, L., Leforestier, J. F., Dimet, J., Doz, F., Chatellier, G., et al. (2010). The readability of information and consent forms in clinical research in France. *PLoS ONE*, 5(5).

Dit artikel beperkt zich ertoe leesbaarheidsformules los te laten op een aantal informed consent-documenten; lezersonderzoek is niet gedaan.

Zoals uit deze voorbeelden blijkt, moesten er heel wat resultaten doorgewerkt worden om te komen tot relevante treffers. Ook in de fase van het beschrijven van aanvankelijk geselecteerde studies vond nog selectie plaats; artikelen die op basis van het abstract en de titel in de kennisbank werden opgenomen, maar bij beter lezen toch niet geschikt bleken te zijn. Bijvoorbeeld omdat alleen oordelen waren verzameld en geen begripsmeting had plaatsgevonden.

4. Hoe zijn de publicaties beschreven?

In de laatste fase hebben wij de geselecteerde artikelen volgens een bepaald model beschreven. Het doel van deze beschrijvingen was: de gebruiker van de kennisbank in staat stellen specifiek te zoeken op literatuur over bepaalde genres, tekstkenmerken of begripseffecten. De meerwaarde van de kennisbank is dus niet alleen de selectie van relevante literatuur over begrijpelijkheid, maar ook het ontsluiten en ordenen daarvan. Dankzij de beschrijvingen

kan de bezoeker van de kennisbank bijvoorbeeld de vraag stellen: welke literatuur is er verschenen over het effect van tekststructuur op het begrip van medische bijsluiters?

We bespreken hieronder kort het beschrijvingsmodel. We merken vooraf nog wel op dat het model niet per publicatie, maar per deelonderzoek wordt toegepast. Heel wat publicaties rapporteren meerdere ‘studies’. Omdat die studies op allerlei kenmerken verschillen, wordt per studie beslist over opname in de kennisbank en wordt elke studie apart beschreven.

4.1 Basisinformatie Het model begint met algemene bibliografische gegevens: auteur, jaartal, volledige titelbeschrijving. Daarnaast is in de kennisbank bij iedere treffer een Google Scholar-icoon te vinden. Wanneer de gebruiker van de kennisbank digitale toegang heeft tot de publicatie, leidt dit icoon daar rechtstreeks heen. Wanneer het artikel meerdere experimenten verslaat, meldt het experimentnummer om welke deelstudie het gaat.

4.2 Samenvatting en conclusie De kennisbank geeft het volledige abstract van het artikel, in dezelfde taal als het artikel zelf. Daarnaast zijn de belangrijkste conclusies nog eens apart weergegeven. Soms zijn die geput uit de samenvatting, maar wanneer die niet voldoet zijn de conclusies uit het artikel zelf gebruikt.

4.3 Genre en genregroep Vervolgens wordt de genregroep gegeven waarin de onderzochte tekst zich bevindt, en het genre. We hebben de volgende domeinen geselecteerd.

1. Gezondheid
2. Financieel
3. Onderwijs
4. Juridisch
5. Wetenschap
6. Argumentatief en persuasief
7. Media
8. Fictie
9. Het gebruik van apparaten
10. Niet-domeingebonden genres
11. Losse tekstfragmenten

Het bleek lastig om de domeinen zo te kiezen dat ze niet overlappen. Neem een geneesmiddelenadvertentie. Die tekst gaat over gezondheid, maar kan verschijnen in de media en daarnaast is hij persuasief van aard. In zulke gevallen is gekozen voor de categorie die het *onderwerp* van de tekst het meest specifiek benoemt, dat is hier dus gezondheid. Evenzo zijn hypotheekvoorwaarden geschaard onder het financiële domein, en niet onder het juridische. Enkele categorieën vergen een nadere toelichting.

- Onderwijsteksten zijn niet teksten over onderwijs, maar teksten die bedoeld zijn voor primair en voortgezet onderwijs. Het standaardvoorbeeld: schoolboekteksten.
- Wetenschappelijke teksten zijn teksten over onderzoek en/of teksten die functioneren in wetenschappelijk onderwijs. Populair-wetenschappelijke teksten worden ook in de groep wetenschap geplaatst. Wanneer het onderzoek werkt met studenten als proefpersonen, is de tekst in principe als populair-wetenschappelijk gecodeerd, behalve wanneer het gaat om een specifiek onderwerp dat een eigen domeinnaam heeft: gezondheid, geld, wetten of apparaten.

Een Kennisbank over Begrijpelijke Taal

- De fictieteksten in deze databank betreffen meestal geen literaire fictie, maar korte verhalen die zijn geconstrueerd door psycholinguïsten ten behoeve van onderzoek.
- Bij niet-domeingebonden teksten gaat het om genres als formulieren, vragenlijsten en websites, die allerlei soorten thema's kunnen betreffen.
- Onder 'losse tekstfragmenten' zijn stimuli geschaard die niet herkenbaar zijn wat betreft genre. Het gaat veelal om de woorden, zinnen en korte tekstjes die worden gebruikt in psycholinguïstisch onderzoek.

In principe plaatsen we elke studie binnen slechts één domein, behalve wanneer daadwerkelijk meerdere domeinen onderzocht zijn in dezelfde studie, bijvoorbeeld zowel persuasieve teksten als verhaaltjes (fictie).

4.4 Boodschapkenmerken en kenmerkgroepen Vervolgens wordt het boodschapkenmerk beschreven waar het onderzoek over rapporteert. Eerst wordt de groep gegeven waarin het kenmerk zich bevindt. In de kennisbank worden 13 groepen onderscheiden:

1. Woordmoeilijkheid
2. Kenmerken op zinsniveau
3. Informatiedichtheid
4. Tekststructuur
5. Lokale markering van tekstsamenhang
6. Globale structuursignalen
7. Tekstvormgeving
8. Kenmerken op tekstniveau
9. Extra informatie
10. Tempo en voordracht
11. Hypertekstkenmerken
12. Gesproken, geschreven of digitaal
13. Visuele elementen

Vervolgens zijn binnen de groepen kenmerken beschreven. Zo is binnen groep 5 (lokale markering van tekstsamenhang) onderscheid gemaakt tussen terugverwijzingen (= referentiele links), verbindingswoorden en signaalzinnen of signaalwoorden. Binnen groep 6 (globale structuursignalen) is onderscheid gemaakt tussen organizers (overzichten van de tekstinhoud), kopjes en benadrukking van tekstgedeeltes.

Bij het coderen op kenmerken moesten lastige keuzes gemaakt worden. Om te beginnen was voorafgaand aan het beschrijven niet voor elke groep duidelijk welke kenmerken nodig waren om de studies in de kennisbank te dekken. Daarom zijn er een aantal studies in de categorie 'overig' terechtgekomen. Er is bijvoorbeeld onderzoek gevonden naar het onthouden van tekstelementen op verschillende niveaus in de tekststructuur: belangrijke elementen staan 'hoog', minder belangrijke 'laag' (Birkmire, 1985; Vipond, 1980). Omdat dit kenmerk niet vooraf was gedefinieerd, zijn deze studies wel geplaatst in de kenmerkgroep 'tekststructuur', maar hebben ze daarbinnen geen eigen kenmerk gekregen.

Naast tekstkenmerken zijn ook bekende adviezen genoemd waarvoor het onderzoek relevant is, om de kennisbank voor praktijkmensen toegankelijker te maken. Voorbeelden van zulke adviezen:

- licht moeilijke woorden toe;
- formuleer concreet en beeldend.

De lijst adviezen is korter dan de lijst kenmerken, omdat er kenmerken zijn waarvoor op dit moment nog geen bekende adviezen bestaan. Daarom is ervoor gekozen om bij het snelzoeken in de kennisbank niet uit te gaan van adviezen maar van kenmerkgroepen, die een volledig beeld bieden van wat de kennisbank te bieden heeft.

4.5 Begripsmetingen en leesprocesmaten We hebben onderscheid gemaakt tussen begripsmetingen en leesprocesmaten. Onder begripsmetingen vatten we allerlei vragen en opdrachten die proefpersonen uitvoeren om te demonstreren wat ze van de tekst begrijpen. Bijvoorbeeld het beantwoorden van gesloten meerkeuzevragen over expliciete tekstinhoud, het maken van cloze-tests of het zoeken naar informatie in de tekst. Procesmetingen richten zich niet op de inhoud maar op het verwerkingsproces, meestal op de snelheid daarvan, soms op de gevolgde route.

Bij de begripsmaten zijn sommige mogelijke onderscheidingen niet gemaakt. Zo was er in sommige artikelen sprake van *immediate recall* en *delayed recall*. Het onthouden van informatie werd meteen getest en bijvoorbeeld 24 uur later nog eens. Dit verschil kwam uiteindelijk niet in de kennisbank.

De meeste van de begripsmaten waren eenduidig. Een uitzondering is de categorie ‘taakuitvoering’. Deze heeft ten eerste betrekking op studies waarin begrip wordt gemeten door instructies daadwerkelijk te laten uitvoeren. Daarnaast is hij gebruikt voor een aantal andere soorten metingen. De belangrijkste daarvan is het *toepassen* van de informatie uit de tekst; daarbij gaat het om het oplossen van problemen met behulp van de tekst. Daarbij kan het gaan om onderwijsteksten waarvan de inhoud moet worden toegepast op een nieuw probleem, jury-instructies die op de gegevens van een rechtszaak moeten worden toegepast of bijsluiters die op hypothetische situaties van geneesmiddelgebruikers moeten worden toegepast. Andere taken die veel minder vaak voorkwamen, waren het uitvoeren van een sorteertaak wanneer die wordt gebruikt als maat voor begrip, het schrijven van een vervolgzin of het verbeteren dan wel becommentariëren van een tekst.

4.6 De aard van de stimuli en de manipulatie Bij elke studie is vermeld of het ging om gesproken, geschreven (papieren) of digitale boodschappen. Verder is de omvang en aard van de stimuli vermeld: ging het om complete teksten, om zinnen, om afzonderlijke woorden, of om visuele elementen? Het onderzoek kan bijvoorbeeld op woordniveau plaatsvinden, zoals in onderzoek naar concrete en abstracte woorden. Ook kan met zinnetjes gewerkt worden, zoals in onderzoek waarin zinnen in de lijdende of bedrijvende vorm aan plaatjes gekoppeld moeten worden.

Verder leek het ons relevant te vermelden hoeveel stimuli (teksten, zinnen, plaatjes enzovoort) er gebruikt zijn in het onderzoek, zodat de generaliseerbaarheid van de onderzoeksresultaten ingeschat kunnen worden. Daarnaast vonden we soms artikelen waarin de stimuli zelf niet te vinden zijn, niet eens voorbeelden daarvan. Het leek ons nuttig om dat te vermelden.

Ten slotte is het belangrijk om te weten in hoeverre de manipulatie enkelvoudig of meervoudig van aard is. Regelmatig worden teksten op meerdere kenmerken tegelijk gemanipuleerd. Zo worden in Steehouder en Jansen (1982) een groot aantal kenmerken van voorlichtingsteksten over huursubsidie tegelijk gevarieerd. Met zulke meervoudige manipulaties kunnen grotere effecten op begrip worden gevonden. Maar voor de gebruiker van de kennisbank is het nuttig om te weten dat zulke effecten niet eenduidig aan afzonderlijke kenmerken kunnen worden toegeschreven. Het onderzoek is veelal praktisch zeer relevant, maar kent beperkingen wat betreft de interne validiteit.

Oorspronkelijk was het de bedoeling dat we ook voor iedere studie zouden vaststellen of de gebruikte tekst ‘realistisch’ was, dat wil zeggen dat de tekst in het dagelijks leven ook op die manier voorkomt. Dat kenmerk is afgefallen omdat het te veel discussie oproep.

4.7 De proefpersonen Over de proefpersonen van het onderzoek wordt het volgende vermeld in de kennisbank: hun leeftijd, opleidingsniveau en hun aantal. Al deze gegevens zijn van belang bij het beoordelen van de generaliseerbaarheid van de onderzoeksresultaten.

4.8 Controlevariabelen en niet-tekstuele onafhankelijke variabelen De interne validiteit van het onderzoek is onder andere afhankelijk van een controle op mogelijk relevante kenmerken van proefpersonen of leestaken. Zulke controlevariabelen worden daarom vermeld. Evenzeer is het belangrijk voor de generaliseerbaarheid van de resultaten als de boodschapvariabelen zijn gekruist met niet-tekstuele onafhankelijke variabelen, waarbij het weer kan gaan om proefpersoonkenmerken of leestaken. Op die manier krijgen we meer inzicht in de robuustheid van het kenmerkeffect, en in eventuele interacties met andere componenten van de leessituatie.

5. Het ontwerp van de kennisbank

De keuzes die hierboven toegelicht zijn, hebben uiteraard direct gevolgen voor de inhoud en het ontwerp van de kennisbank. Daarom zijn het zoeken naar literatuur en het overleg over het ontwerp van de kennisbank gelijk opgegaan. Het ontwerpproces van de kennisbank bestond uit twee fasen. In de eerste fase is in samenwerking met de afdeling Informatiekunde van de VU een programma van eisen samengesteld. Daarna is in samenwerking met softwarebureau *Iblux* een ontwerp gemaakt dat vervolgens is gevuld met inhoud en een aantal keren is aangepast.

5.1 Fase 1: Programma van eisen In de eerste fase hebben we ons georiënteerd op bestaande kennisbanken zoals die op het Internet te vinden zijn. Bovendien hebben we een globale analyse gemaakt van doelen en doelgroepen van de kennisbank. In deze fase zijn de volgende beslissingen genomen.

- Doel: met de kennisbank brengen we de wetenschappelijke literatuur in kaart die de afgelopen decennia is verschenen over kenmerken van teksten die de begrijpelijkheid beïnvloeden. Deze literatuur wordt beschikbaar gesteld via een titelbeschrijving en een samenvatting. Bovendien biedt de kennisbank de mogelijkheid om specifiek te zoeken naar afzonderlijke genres, tekstkenmerken, begripseffecten en combinaties daarvan.
- Taal: aangezien veel literatuur in internationale tijdschriften is verschenen, zou de ken-

nisbank ook voor een internationaal publiek toegankelijk moeten zijn. Wij kennen ook geen voorbeeld van zo'n kennisbank in het Engelse taalgebied. Daarom is er naast de Nederlandse ook een Engelstalige versie van de kennisbank: *kennisbank-begrijpelijketaal.nl/en*. Zo is er dan ook een internationale kennisbank met publicaties in diverse Europese talen.

- Doelgroep: we onderscheiden een academische en een professionele doelgroep. Dat zijn dus enerzijds studenten, docenten en onderzoekers en anderzijds tekstschrijvers, adviseurs en managers op het terrein van tekstkwaliteit.
- Voorkennis: we veronderstellen voorkennis op zowel methodisch als taalkundig terrein. Dit betekent dat we gebruik zullen maken van termen als *type-token ratio*, ook al zal niet elke gebruiker van de kennisbank met die termen vertrouwd zijn. Voor de minder deskundige gebruikers is een termenlijst gemaakt die uitleg biedt. Idealiter zouden de toelichtingen vanzelf moeten verschijnen wanneer de muis over de term gaat, maar dat bleek technisch buiten ons budget te vallen.
- Manieren van zoeken: we kiezen voor een snelkeuzemenu en een optie voor geavanceerd zoeken, waarmee verfijndere zoekacties mogelijk zijn. Met het snelkeuzemenu kan gezocht worden op *genregroep*, *kenmerkgroep*, en *begripsmaat*. Met geavanceerd zoeken ontstaat de mogelijkheid om te zoeken op combinaties van waarden uit het merendeel van de geannoteerde velden. Een tweede vorm van verfijnd zoeken wordt mogelijk gemaakt in de lijst met resultaten: na een eerste zoekactie kan de gebruiker met behulp van de knop *verfijn zoekactie* binnen deze eerste selectie dieper zoeken naar specifieke literatuur.
- Gesloten en open velden: Alle zoekvelden bieden een meerkeuzemenu, met uitzondering van de velden voor de auteur, de titel en samenvatting. Daarin kunnen zelfgekozen woorden ingevoerd worden.
- Zoeken in de output: er worden geen nadere functies ontwikkeld om binnen een lijst met treffers te zoeken naar specifieke termen, anders dan de hierboven genoemde verfijningsoptie.
- Combineren van zoektermen: zoektermen uit verschillende velden kunnen alleen worden gecombineerd door *EN*. De combinatie van *woordmoeilijkheid* en *gezondheid* levert studies op over woordeffecten in gezondheidsteksten. Als daarentegen waarden binnen eenzelfde variabele gecombineerd worden, wordt die combinatie als *A OF B* geïnterpreteerd. Dus *woordmoeilijkheid* levert samen met *tekstvormgeving* treffers op die over een van beide kenmerken (of allebei) gaan.
- De literatuur die vrij toegankelijk is via het Internet, zoals Google Scholar, wordt via een apart logo aanklikbaar gemaakt voor de bezoeker.

Nadat deze keuzes gemaakt waren, is *Iblux* aan het werk gegaan om een concept van de kennisbank te ontwikkelen.

Discussies over het programma van eisen werden mede gevoed door de resultaten van de eerste verkenningen van de literatuur. Met vakreferenten van de universiteitsbibliotheek is overleg gevoerd over bestaande systemen van annotatie en over zoekmethoden. Tijdens het annoteren van publicaties moesten er talloze keuzes gemaakt worden die gevolgen hadden voor de uitwerking van de kennisbank. Toen een concept van de interface eenmaal vaststond, konden projectmedewerkers aan de slag gaan met het annoteren van de publicaties. Twijfelgevallen werden wekelijks in de projectgroep besproken. Soms leidde dat tot een nadere

uitwerking van een ontwerpkeuze of tot een herziening van een eerder gemaakte keuze.

5.2 Fase 2: Ontwerp van de Kennisbank Een eerste concept van de kennisbank is begin december 2010 gepresenteerd. In de periode daarna zijn er enkele wijzigingen doorgevoerd. Het fotomateriaal is meer toegespitst op talige associaties. In het menu is een knop toegevoegd waaronder overzichtsartikelen gepresenteerd worden. Het eindresultaat is hieronder te zien (zie Figuur 1).



Figuur 1: Startpagina van de Kennisbank Begrijpelijke Taal

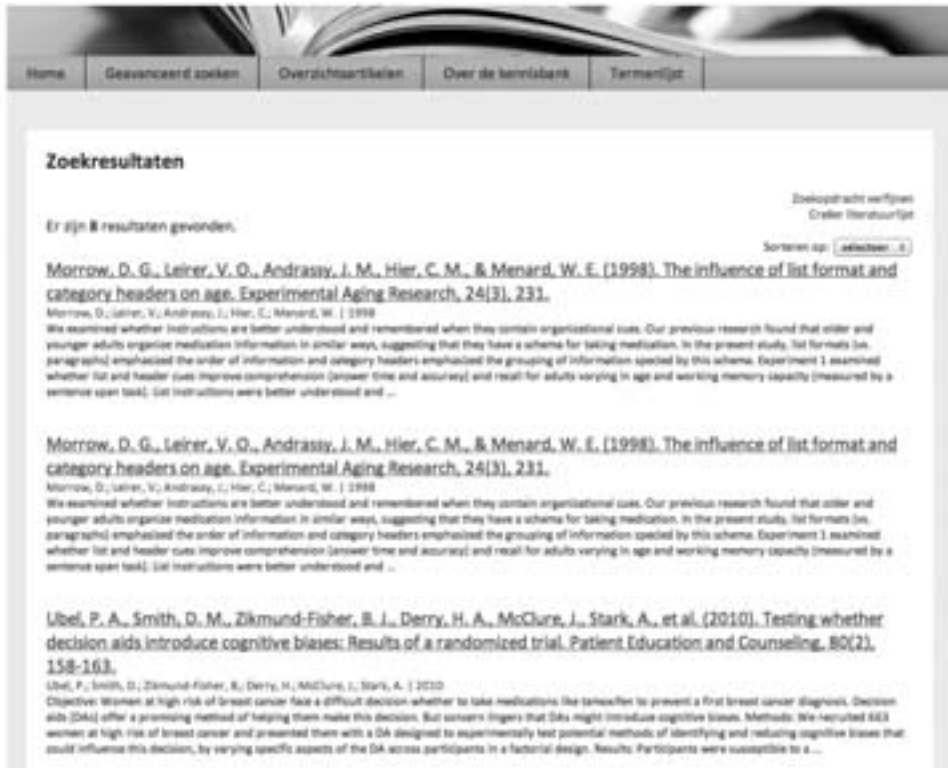
In het scherm *Geavanceerd zoeken* kan de bezoeker gebruik maken van maar liefst 16 kenmerken: auteurs, titel, samenvatting, publicatiejaar, leeftijd van de proefpersonen, genregroep, genre, tekstkenmerksgroep, tekstkenmerken, het taaladvies waarvoor een studie relevant is, begripmaten, leesprocesmaten, niet-tekstuele onafhankelijke variabelen (bv. opleidingsniveau), controlevariabelen (bv. intelligentie), modaliteit (mondeling, papier of digitaal) en stimulus (teksten, woorden, zinnen, visuele elementen). Via combineren ontstaan talloze mogelijkheden.

Een enkele keer is een keuzemenu te lang voor de gebruiker: in het beschrijvingsmodel annoteerden we de artikelen met behulp van een lijst met 57 tekstkenmerken. De gebruiker krijgt echter alleen de tekstkenmerken te zien binnen de kenmerksgroep die hij aanklikt. Dus pas na het kiezen van 'woordmoeilijkheid' verschijnen daarnaast kenmerken als woordlengte, woordfrequentie, woordconcreetheid e.d.

Een voorbeeld. Wie in het domein *gezondheid* zoekt naar effecten van *tekststructuur* op de begrijpelijkheid van *geneesmiddeleninformatie*, krijgt allereerst te zien dat deze zoekactie acht treffers heeft opgeleverd (zie Figuur 2). Vervolgens blijkt dat de eerste twee treffers verwijzen

naar dezelfde publicatie. In het artikel van Morrow e.a. worden namelijk twee afzonderlijke studies gerapporteerd over effecten van structuurkenmerken. Rechtsboven worden drie opties aangeboden:

- orden de treffers op een andere manier (jaartal of auteur)
- verfijn de zoekopdracht
- maak een bibliografie.



Figuur 2: De eerste drie (van de acht) treffers bij geavanceerd zoeken op: gezondheid & geneesmiddeleninformatie & tekststructuur

Met de knop *Creër literatuurlijst* komt de gebruiker in een nieuw scherm, waarin relevante publicaties geselecteerd kunnen worden. Na die selectie levert de knop *Exporteer* een Word-bestand op met daarin een literatuurlijst waar de geselecteerde titels in zijn opgenomen. Daarnaast is er een optie toegevoegd om in het scherm met resultaten een verfijnde selectie uit te voeren.

Wie in de output van de kennisbank op de eerste treffer klikt, komt terecht in een scherm waarin alle gegevens over deze treffer te vinden zijn (zie Figuur 3). Allereerst zien we de bibliografische gegevens, dan een link naar Google Scholar en daaronder de samenvatting (in de taal van de publicatie).



Figuur 3: Scherm met informatie over een treffer (deel 1)

En daar weer onder staan alle andere gegevens (zie figuur 4). Allereerst is aangegeven voor welk schrijfadvisies het onderzoek relevant is, vervolgens zien we de begripsmaten, procesmaten en de conclusies. Tot slot volgen de details wat betreft het materiaal, de manipulaties en de proefpersonen.



Figuur 4: Scherm met informatie over een treffer (deel 2)

6. Een kijkje in de Kennisbank

Wat zit er in de Kennisbank Begrijpelijke Taal? Er zijn 459 artikelen gevonden die empirisch onderzoek rapporteren naar de effecten van boodschapkenmerken op begrijpelijkheid. Menig artikel rapporteert meer dan één studie. In totaal vonden we 693 studies. De meeste studies zijn verricht in de volgende domeinen:

- wetenschap (168);
- onderwijs (120);
- gezondheid (119).

Bij onderwijsteksten gaat het grotendeels om onderzoek waarbij leerlingen van scholen gemanipuleerde teksten voorgelegd kregen. Die teksten konden over allerlei mogelijke onderwerpen gaan. Bij wetenschap vinden we twee soorten teksten. Ten eerste zijn er teksten over allerlei studiegerelateerde onderwerpen die aan studenten aan universiteit zijn voorgelegd (bv. over psychologische of geologische onderwerpen), ten tweede populair-wetenschappelijke teksten. Duidelijker qua thematiek is de groep gezondheid, die altijd gaat over gezondheidszorg-onderwerpen.

Verrassend is het geringe aantal studies dat we vonden in het domein van de financiële wereld: slechts zes studies komen uit dit domein. Iets meer studies vonden we in het juridische domein (16). Hoewel financiële en juridische teksten door leken veelal bijzonder moeilijk worden geacht, is er dus niet veel onderzoek naar gedaan.

Naar welke boodschapkenmerken is vooral onderzoek verricht? Hieronder staan de voornaamste kenmerken in volgorde van intensief naar nauwelijks onderzocht. We maken een onderscheid tussen het onderzoek naar visuele en naar talige kenmerken. Er zijn ruim 199 studies naar effecten van visuele kenmerken, die als volgt zijn onderverdeeld naar typen visualisaties.

- Illustraties 97
- Animaties 36
- Schema's 48
- Iconen 26
- Grafieken 21
- Tabellen 21

De meeste aandacht gaat uit naar illustraties. Veel onderzoek naar effecten van visualisaties is gedaan in de context van educatief materiaal, iets minder is het domein van gezondheidsinformatie onderzocht. Maar ook in de context van instructies bij apparaten is aandacht besteed aan effecten van visualisaties.

De meeste studies in de kennisbank gaan over effecten van talige kenmerken. De globale structuur van teksten krijgt daarin de meeste aandacht (154 studies). Het gaat dan bijvoorbeeld om effecten van kopjes (79) en zogenaamde organizers (57), dat wil zeggen globale vooruitblikken op de inhoud en opbouw van de tekst. Ook is er veel aandacht voor de tekststructuur zelf (135). Daarbij gaat het opvallend vaak om de volgorde van tekstonderdelen (73).

- Globale structuursignalen 154
- Tekststructuur 135

Een Kennisbank over Begrijpelijke Taal

• Extra informatie	100
• Zinskenmerken	94
• Lokale markering tekstsamenhang	87
• Woordmoeilijkheid	84
• Kenmerken op tekstniveau	70
• Vormgeving van tekst	59
• Gesproken, geschreven of digitaal	37
• Hypertekst kenmerken	21
• Tempo en voordracht	13

In de middengroep vinden we studies naar effecten van woordmoeilijkheid en kenmerken van de zin, zoals de zinslengte en grammaticale kenmerken van de zin (bijzinnen, tangconstructies en de lijdende vorm). Ook vinden we daar publicaties over het toevoegen van extra informatie (zoals voorbeelden) en over de lokale markering van tekststructuur (relaties tussen elementen binnen passages, zoals verbindingswoorden en terugverwijzingen). Kenmerken op tekstniveau zijn bijvoorbeeld het perspectief van de tekst en globale maten voor tekstcomplexiteit.

Wat minder onderzoek is er in de kennisbank te vinden over mediumkeuze (maakt het uit of een tekst voorgelezen wordt of zelf gelezen wordt?) en kenmerken van hypertekst (zoals de formulering van linklabels). Weinig onderzoek is gedaan naar het effect van kenmerken van mondelinge presentaties, zoals tempo en voordracht.

Naast de talige elementen is er aandacht voor vormgeving en voor visuele elementen. Bij de vormgeving van de tekst (59 studies) gaat het om het formaat van opsommingen en bijvoorbeeld regellengte. De studies over visuele elementen (199 in getal) gaan bijvoorbeeld over:

• Illustraties	96
• Schema's	43
• Animaties	36
• Iconen of pictogrammen	26
• Grafieken	21
• Tabellen	19

Wat zijn populaire methoden om begrip te meten? Het meest gebruikt wordt de meerkeuzevraag, onmiddellijk gevolgd door de open vraag. Er wordt vaker naar expliciete dan naar impliciete informatie gevraagd; regelmatig wordt zowel naar expliciete als impliciete informatie gevraagd. Daarom is hieronder het totaal van de studies met expliciete open vragen en die met impliciete open vragen ook hoger dan het aantal studies met open vragen als geheel. De hekkensluis is de cloze-test.

• Meerkeuzevragen totaal	225
o Meerkeuzevragen naar expliciete informatie	208
o Meerkeuzevragen naar impliciete informatie	73
• Open vragen totaal	214
o Open vragen naar expliciete informatie	160
o Open vragen naar impliciete informatie	118

• Vrije reproductie	158
• Herkenning van de stimulus of gerelateerd materiaal	114
• Taakuitvoering	78
• Gestuurde reproductie	76
• Samenvatting	38
• Het vinden van informatie	19
• Cloze-test	18

We lichten een paar methodenamen toe. Bij *herkenning* is sprake van een test waarbij elementen die letterlijk in de tekst staan of direct aan de tekst gerelateerd zijn aan de proefpersoon gepresenteerd worden. Proefpersonen moeten dan melden of zij dit element wel of niet gezien hebben, of een lexicale decisie maken. Bij *gestuurde reproductie* wordt bijvoorbeeld een schema aangeboden waarin de proefpersoon elementen moet plaatsen, bijvoorbeeld in termen van oorzaak en gevolg of in de juiste chronologische volgorde. Bij de *cloze-test* wordt een tekst aangeboden met systematisch weggelaten woorden die de proefpersoon moet invullen.

Met de bovengenoemde methoden wordt het *product* van een leesactiviteit bepaald. Dat product is immers een bepaalde mentale representatie van de gelezen tekst. Daarnaast wordt veelal ook gekeken naar het *proces* tijdens het lezen. Hoe wordt dat vastgesteld? We vonden 342 studies waarin ook uitspraken over het proces gedaan worden. De meest populaire procesmaat is de *leestijd* (219 studies), gevolgd door de *antwoordtijd* (134). Veel minder is gekeken naar de leesvolgorde van de proefpersonen (38). Daarbij gaat het soms om lokale maten zoals het aantal regressies naar eerdere woorden in de zin, maar vaker om meer globale routekenmerken zoals het navigatiepad in een hypertext of het schakelen tussen tekst en plaatjes.

Ten slotte kunnen we dankzij de kennisbank inzicht krijgen in een andere factor die eveneens van grote invloed is op het begrip van teksten: de kenmerken van de lezer. Er zijn 190 studies waarin lezerkenmerken aan de orde waren als *onafhankelijke* variabelen. Dit betekent dat de onderzoekers erop uit waren iets te weten te komen over effecten van de lezerkenmerken op de begrijpelijkheid van teksten in combinatie met bepaalde tekstkenmerken. Het meest populaire – en misschien ook wel belangrijkste – kenmerk is de *voorkennis* van de lezer. De minst onderzochte kenmerken zijn intelligentie, geslacht en leerstijl. Hieronder staan de verschillende kenmerken in volgorde van populariteit.

• Voorkennis	83
• Leeftijd	52
• Leesvaardigheid	45
• Werkgeheugen	33
• Opleiding	15
• Geslacht	8
• Intelligentie	8
• Leerstijl	5

Wie grasduint in de kennisbank, zal ontdekken dat er nog veel meer mogelijkheden zijn om te verkennen in welke richting het begrijpelijkheidsonderzoek zich beweegt. In nog te publiceren artikelen zullen we een aantal van zulke proefboringen uitvoeren, onder andere gericht op bepaalde genres, zoals medische informatie voor patiënten, en op bepaalde tekstkenmerken zoals structuurmarkering.

Tot zover hebben we de inhoud van de kennisbank vrij globaal weergegeven. We sluiten deze bespreking af met een voorbeeld van een gevalstudie over een concreet probleem. De vraag is: wat weten we over het effect van de lijdende vorm? Als we op dat tekstkenmerk zoeken, vinden we 14 studies; overigens zijn er heel wat publicaties van vóór 1980 die daarmee buiten beschouwing blijven.

Van de 14 jongere publicaties blijken er 6 een meervoudige manipulatie te gebruiken: Charrow & Charrow (1979), Handel e.a. (2001), Kang e.a. (2009), Leroy e.a. (2010), Ulijn & Strother (1990) en Wenger & Spyridakis (1993). Hoewel het gaat om interessante tekstverbeteringstudies, leveren ze te weinig toegespitste informatie op over de lijdende vorm om verder te bespreken. Bijvoorbeeld: Ulijn en Strother (1990) hebben van 10 zinnen de zinsbouw vereenvoudigd, waarbij in twee zinnen de lijdende vorm samen met andere kenmerken is gemanipuleerd, en slechts in één zin alleen de lijdende vorm op zich is veranderd. In hun begripsmeting is het effect van deze zinnen niet meer te isoleren van andere ingrepen. En in Kang e.a. (2009) zien we naast de lijdende vorm ingrepen in het gebruik van medische vaktermen, het lettertype en in witregels. De studies met enkelvoudige manipulaties hebben we bij elkaar gezet in Tabel 1.

Tabel 1: Korte signalen van studies met enkelvoudige passiefmanipulaties

Referentie	Genre	Begripsmeting	Resultaat en commentaar
<i>Bostian (1983)</i>	Nieuwsberichten	Leessnelheid (= aantal zinnen dat in een bepaalde tijd gelezen werd); begripsvragen naar expliciete informatie	In de passieve tekst werden minder zinnen gelezen dan in de actieve tekst; er was geen verschil in begrip.
<i>Carrithers (1989)</i>	Losse zinnen	Leestijden bij woord-voorwoord lezen	Passieve zinnen met 'normale' werkwoorden worden verrassend genoeg sneller gelezen; ze worden wel langzamer gelezen met 'ergatieve' werkwoorden zoals <i>amaze</i> .
<i>Dabrowska & Street (2006)</i>	Losse zinnen	Gevraagd werd wie de actor was in de beschreven gebeurtenis.	Hoogopgeleide sprekers presteren altijd op plafondniveau. Laagopgeleide sprekers presteren slechter op de passieve zinnen dan op de actieve variant, wanneer de zin gaat over onwaarschijnlijke gebeurtenissen.
<i>Ferreira (2003). Experiment 1</i>	Losse zinnen	Gevraagd werd wie de actor was in de beschreven gebeurtenis.	Bij onwaarschijnlijke gebeurtenissen leveren passieve zinnen meer foute antwoorden op dan de actieve variant daarvan.
<i>Ferreira (2003). Experiment 2.</i>	Losse zinnen	Gevraagd werd wie de actor was in de beschreven gebeurtenis.	Bij onwaarschijnlijke gebeurtenissen leveren passieve zinnen meer foute antwoorden op dan corresponderende gekloofde zinnen met het onderwerp voorop.
<i>Lenzner e.a. (2010)</i>	Vragen in vragenlijst	De tijd die nodig was om de vraag te beantwoorden; het aantal 'geen-mening' antwoorden.	Twee vragen in de passieve variant kosten meer tijd dan de actieve tegenhangers; er was geen verschil in de antwoorden zelf
<i>Street & Dabrowska (2010). Experiment 1</i>	Losse zinnen	Het juiste plaatje kiezen bij een zin.	Bij laag opgeleide sprekers leveren passieve zinnen meer fouten op dan actieve.
<i>Street & Dabrowska (2010). Experiment 2</i>	Losse zinnen	Het juiste plaatje kiezen bij een zin.	Na een training maken laag opgeleide sprekers bij passieve zinnen vrijwel geen fouten meer.

Bostian (1983) en Lenzner e.a. (2010) zijn de enige studies die een passief-manipulatie onderzochten in een authentieke tekst. Het onderzoek van Bostian lijdt helaas aan gebreken. Zo is er een alternatieve verklaring voor het feit dat in de passieve tekst minder zinnen gelezen werden: Bostian meldt namelijk dat de zinnen in die tekst langer waren dan in de actieve tekst (15.7 versus 13.5 woorden). Interessant aan de studie van Lenzner e.a. is de bijzondere uitkomstvariabele: de tijd die nodig is om een surveyvraag in de bedrijvende dan wel lijdende vorm te beantwoorden. Voor de passiefmanipulatie gaat het in zijn onderzoek overigens slechts om twee vragen, zodat de generaliseerbaarheid nog niet vaststaat.

Verder wordt het lijstje gedomineerd door psycholinguïstisch onderzoek op basis van afzonderlijke zinnen. In Carrithers (1989) wordt aangenomen dat passiefzinnen complexer zijn, maar wordt betwijfeld of complexiteit moet lijden tot langere leestijden. Passiefzinnen blijken niet langzamer te worden gelezen, behalve wanneer er ergatieve werkwoorden in staan.

Natuurlijk is leessnelheid slechts een indirecte maat voor begripsproblemen. In dat opzicht zijn de studies van Ferreira, Street en Dabrowska interessanter. Zij stellen vast dat bij de interpretatie van passieve zinnen fouten gemaakt worden, met name bij onwaarschijnlijke gebeurtenissen en alleen door lager opgeleide sprekers. Bij Street en Dabrowska (2010) blijkt verder dat er na training veel minder fouten gemaakt worden, wat erop wijst dat het probleem met de lijdende vorm niet zozeer ligt in beperkingen van het werkgeheugen (zoals met tangconstructies het geval is), maar in een gebrek aan ervaring met deze constructie bij lager opgeleiden.

7. De kennisbank in het academisch onderwijs

Een exercitie zoals hierboven beschreven, leidt tot een wetenschappelijk verantwoorde samenvatting van relevant onderzoek op een specifiek tekstkenmerk. Met de kennisbank zijn erg veel vergelijkbare exercities mogelijk. Dat is niet alleen interessant voor de onderzoeker of de geïnteresseerde professional, maar ook voor docenten en studenten in het universitaire onderwijs. Wij hebben inmiddels ervaring opgedaan met het gebruik van de kennisbank in ons eigen BA-onderwijs. Studenten werkten in het kader van hun BA-eindwerkstuk aan een overzichtsartikel voor een specifieke vraag met behulp van literatuur uit de kennisbank. Daarbij kregen ze de opdracht om niet alleen de voor die vraag relevante literatuur te selecteren en te beschrijven, maar ook op zoek te gaan naar mogelijke aanvullingen voor de kennisbank. Dit bleek een zeer leerzame afsluiting te zijn van hun BA-opleiding. Studenten moesten zelfstandig een flinke verzameling van ruim twintig studies bestuderen, waarbij zij een beroep moesten doen op hun theoretische, methodische en statistische kennis van empirisch onderzoek naar taal en communicatie. In hun bespreking moesten zij bovendien over de grenzen van de verschillende studies heen kijken. Zij maakten vergelijkingen tussen verschillende manieren waarop begripelijkheid getoetst was, tussen verschillende theoretische benaderingen waarop de studies een beroep deden, en tussen de uitkomsten die de ene keer wel en de andere keer geen of een tegengesteld effect lieten zien van een bepaalde ingreep. En de presentatie van al die kennis in een overzichtelijk, goed leesbaar, samenvattend artikel vergde veel van hun schrijfvvaardigheid.

In andere cursussen zullen we ook op kleinere schaal gebruik maken van de kennisbank, door studenten te vragen presentaties te verzorgen over kleinere thema's, in nota's gebruik te maken van relevante literatuur en in scripties te tonen dat zij verder zijn gegaan dan een korte verkenning van Google bij hun zoektocht naar literatuur.

8. Hoe verder met de kennisbank?

De *Kennisbank Begrijpelijke Taal* is naar ons weten de eerste database waarmee onderzoekers, studenten en professionals zicht krijgen op het empirisch onderzoek naar begrijpelijkheid. Is zo'n kennisbank wel echt nodig? We menen van wel. Het digitale tijdperk wekt weliswaar de suggestie dat alle publicaties voor het grijpen liggen, maar onze ervaring is anders: er is zowel overvloed als schaarste. Een combinatie van zinnige zoektermen kan een oogst van potentieel interessante publicaties opleveren. Maar die oogst kent beperkingen. Ten eerste moeten vaak verschillende bestanden worden gecombineerd om zeker te zijn dat er geen belangrijke tijdschriften ontbreken. Ten tweede is onze indruk dat boeken er in veel zoeksystemen bekaaid afkomen. De beperkte tijd voor dit project maakte het onmogelijk om die indruk nader te onderbouwen. Ten derde, en dat is wellicht in de praktijk het belangrijkste, moeten er heel wat potentiële treffers worden doorgenomen om tot echt relevante treffers te komen. Het is met name hier dat de kennisbank winst kan opleveren. Om deze winst te consolideren, en Nederland tot een internationaal expertisecentrum van begrijpelijkheid te maken, zijn vervolginvesteringen naar onze mening nodig.

8.1 Update van de kennisbank Het NWO-programma, dat dit project mogelijk heeft gemaakt, zal een looptijd van ongeveer vijf jaar hebben. Het is zeer wenselijk in die vijf jaar de kennisbank periodiek bij te werken, door recente publicaties toe te voegen. Daarmee worden twee doelen gediend. Ten eerste blijft de kennisbank zo bij de tijd. Ten tweede kunnen we in die periode een realistische schatting maken van de tijd die nodig is voor periodieke updates. Dan kan financiering gezocht worden voor onderhoud op de langere termijn.

Aanvullingen verwachten we op twee manieren. Ten eerste zullen collega-onderzoekers ons hopelijk attenderen op kandidaat-studies voor de kennisbank. De opgestuurde studies moeten worden beoordeeld en geannoteerd als zij een plaats in de kennisbank verdienen. Ten tweede zullen we langs digitale weg de kennisbank bij de tijd proberen te houden. Daarvoor staan verschillende methoden ter beschikking.

- Ten eerste kunnen voor elk deelgebied voor de meest geciteerde studies nieuwe citaties worden bijgehouden.
- Ten tweede zullen we voor een reeks trefwoorden bij de bibliotheek een periodieke 'alert' aanvragen, waarmee de nieuwe publicaties binnen het omschreven terrein automatisch worden opgestuurd.
- Ten derde is het mogelijk de complete zoektermenverzameling opnieuw te gebruiken in een search waarin alleen publicaties van 2010 en later worden opgevraagd.

We zullen nagaan welk van deze methoden het meest succesrijk is en tegelijkertijd het meest haalbaar, gezien de beschikbare tijd.

8.2 Test van de kennisbank In het ontwikkelproces van de kennisbank was geen tijd en geen geld begroot voor een pretest van de website. Te voorspellen valt dat reacties van gebruikers zullen leiden tot nieuwe wensen. Misschien is er vervolgens een nieuwe investering nodig in verdere ontwikkeling van de applicatie. De projectgroep zal wensen inventariseren en indien nodig een nieuwe projectaanvraag opstellen voor eventuele verdere ontwikkeling.

9. Conclusie

De Kennisbank Begrijpelijke Taal biedt professionals, wetenschappers en studenten een grote hoeveelheid onderzoeksliteratuur over een moeilijk toegankelijk studiegebied. De literatuur is namelijk verspreid over verschillende disciplines en een grote hoeveelheid bronnen. Geen enkele zoekmachine biedt met een eenvoudige reeks zoektermen een overzicht zoals nu beschikbaar is met de kennisbank.

Niet alleen zijn deze publicaties nu bijeengebracht in de kennisbank, ze zijn ook geannoteerd. Daarmee kan de bezoeker, ook zonder het artikel zelf te raadplegen, kennis nemen van allerlei specifieke gegevens zoals de hoeveelheid proefpersonen, de onderzoeksopzet en de conclusies. Ten slotte zijn veel van de publicaties direct op te vragen via de kennisbank, hoewel dat vanuit universiteiten en onderzoekscentra eenvoudiger zal zijn dan voor de particulier, omdat de digitale toegang vaak geregeld is via abonnementen bij uitgevers.

Voor het vakgebied van de Taalbeheersing, of in ruimere zin, het gebied van tekstverwerking en tekstontwerp, is dit initiatief uniek. Wij kennen althans geen vergelijkbaar overzicht. Ook in het professionele veld van tekstschrijvers en communicatiespecialisten is dit niet beschikbaar. Daarmee is dit resultaat van het NWO-project *Begrijpelijke Taal* een bijdrage aan zowel de verdere professionalisering van het vakgebied als aan de wetenschappelijke bestudering van factoren die teksten meer of minder begrijpelijk maken.

Noten

- 1 Zie: http://www.nwo.nl/nwohome.nsf/pages/NWOA_7U6AZN
- 2 Aan de Kennisbank Begrijpelijke Taal werkten mee: Ellen Bosch, Charlotte van Hooijdonk, Leo Lentz, Henk Pander Maat en Anne-Wil van Wijlen. De vormgeving is verzorgd door Erik Vuuregge (Iblux). Rogier Kraf verzorgde een afzonderlijk onderdeel over tools voor begrijpelijkheid. De kennisbank en deze publicatie zijn mogelijk gemaakt met steun van het NWO-programma Begrijpelijke Taal (project. Nr. 321-70-050).

Bibliografie

- Anderson, R. C., & Davison, A. (1988). Conceptual and Empirical Bases of Readability Formulas. In A. Davison & G. M. Green, *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 23-53). Hillsdale, NJ: Lawrence Erlbaum.
- Birkmire, D. P. (1985). Text processing: The influence of text structure, background knowledge, and purpose. *Reading Research Quarterly*, 20(3), 314-326.
- Bostian, L. R. (1983). How active, passive and nominal styles affect readability of science writing. *Journalism Quarterly*, 60(4), 635-40.

- Carrithers, C. (1989). Syntactic complexity does not necessarily make sentences harder to understand. *Journal of Psycholinguistic Research*, 18(1), 75-88.
- Charrow, R. P., & Charrow, V. R. (1979). Making legal language understandable: A psycholinguistic study of jury instructions. *Columbia Law Review*, 79(7), 1306-1374.
- Dabrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, 28(6), 604-615.
- Diamond, S.S., & Levi, J.L. (1996). Improving decisions on death by revising and testing jury instructions. *Judicature*, 79(5), 224-232.
- Douma, W.H. (1960). De leesbaarheid van landbouwbladen: een onderzoek naar en een toepassing van leesbaarheidsformules. Landbouwhogeschool Wageningen, Afdeling Sociologie en Sociografie, Bulletin nr. 17.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164-203.
- Fuchs, J., & Hippius, M. (2007). Inappropriate dosage instructions in package inserts. *Patient Education and Counseling*, 67(1-2), 157-168.
- Handel, D., Fluck, H., Forster, M., & Blaha, M. (2001). Burgernahe Verwaltungstexte: ein Kooperationsprojekt der Stadt Bochum und der Ruhr-Universität zur Textoptimierung. *Fachsprache*, 23(3-4), 139-152.
- Hancock, H.E., Fisk, A.D., & Rogers, W.A. (2005). Comprehending product warning information: age-related effects and the roles of memory, inferencing and knowledge. *Human Factors*, 47(2), 219-234.
- Kang, E. Y., Fields, H. W., Kiyak, A., Beck, F. M., & Firestone, A. R. (2009). Informed consent recall and comprehension in orthodontics: Traditional vs improved readability and processability methods. *American Journal of Orthodontics and Dentofacial Orthopedics*, 136(4), 488.e1-488.e13.
- Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L.G. Nilsson (Red.), *Perspectives on Memory Research* (pp. 329-366). Hillsdale, New Jersey: Erlbaum.
- Lentz, L. e.a. (2000). Tekstschrijver als beroep: een nieuwe peiling. *Tekst[blad]*, 6(3), 33-38.
- Lenzner, T., Kaczmarek, L., & Lenzner, A. (2010). Cognitive Burden of Survey Questions and Response Times: A Psycholinguistic Experiment. *Applied Cognitive Psychology*, 24(7), 1003-1020.
- Leroy, G., Helmreich, S., & Cowie, J. R. (2010). The influence of text characteristics on perceived and actual difficulty of health information. *International Journal of Medical Informatics*, 79(6), 438-449.
- Meyer, B. J., Marsiske, M., & Willis, S. L. (1993). Text processing variables predict the readability of everyday documents read by older adults. *Reading Research Quarterly*, 28(3), 234-249.
- Moxey, L. M., Sanford, A. J., Sturt, P., & Morrow, L. I. (2004). Constraints on the formation of plural reference objects: The influence of role, conjunction, and type of description. *Journal of Memory and Language*, 51(3), 346-364.
- Murray, J. D. (1995). Logical connectives and local coherence. In R. F. J. Lorch, & E. J. O'Brien (Red.), *Sources of coherence in reading* (pp. 107-125). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc; England.
- Murray, J. D. (1997). Connectives and narrative text: The role of continuity. *Memory & Cognition*, 25(2), 227-236.
- Murray, N. M., Manrai, L. A., & Manrai, A. K. (1993). Public policy relating to consumer comprehension of television commercials: A review and some empirical results. *Journal of Consumer Policy*, 16(2), 145-170.
- O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, 43(2), 121-152.
- Pander Maat, H., Boer, N. de, & Timmermans, C. (2009). *De gebruiksvriendelijkheid van hypotheekinformatie: een lezersonderzoek*. Rapport van een onderzoek in opdracht van de Autoriteit Financiële Markten. Universiteit Utrecht.
- Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, 92(1), 85-95.
- Salmeron, L., Kintsch, W., & Kintsch, E. (2010). Self-regulation and link selection strategies in hypertext. *Discourse Processes*, 47(3), 175-211.

- Schraw, G. (1998). Processing and recall differences among seductive details. *Journal of Educational Psychology*, 90(1), 3-12.
- Steehouder, M., & Jansen, C. J. M. (1982). De effectiviteit van voorlichtingsteksten. *Tijdschrift voor Taalbeheersing*, 4(4) 293-313.
- Street, J. A., & Dabrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua*, 120(8), 2080-2094.
- Ulijn, J. M., & Strother, J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading*, 13(1), 38-54.
- Vipond, D. (1980). Micro- and macroprocesses in text comprehension. *Journal of Verbal Learning & Verbal Behavior*, 19(3), 276-296.
- Wenger, M. J., & Spyridakis, J. H. (1993). Reduced text structure at two text levels: Impacts on the performance of technical readers. *Journal of Technical Writing and Communication*, 23(4), 333-352.