

ReferentieBestand Belgisch-Nederlands
RBBN1.0
gebruikshandleiding

Juni 2005

I. Algemeen	3
Totstandkoming	3
Karakterisering	3
Doel	3
II. Toelichting bij de database	7
Datacategorieën: veldnamen en -waarden	7

I. Algemeen

Totstandkoming

- opdrachtgever: NTU , als initiatief van de voormalige CLVV
- begeleiders: Prof Dr. Willy Martin (Vrije Universiteit Amsterdam) en Prof. Dr. Willy Smedts (KU Leuven).
- uitvoerders: lic. T. Colleman, lic. S. Dewaert en lic. V. van der Voort
- periode: 1998/9 en 2004/5
- financiering: Periode I : CLVV-NTU
Periode II : TST-centrale (INL-NTU)

Karakterisering

Het bestand bevat ca. 4000 belgicismen die zijn gesubcategoriseerd volgens een systeem dat door W. Martin in 1997 is ontworpen en door K. Deschamps en W. Smedts in 2002 is gevalideerd.

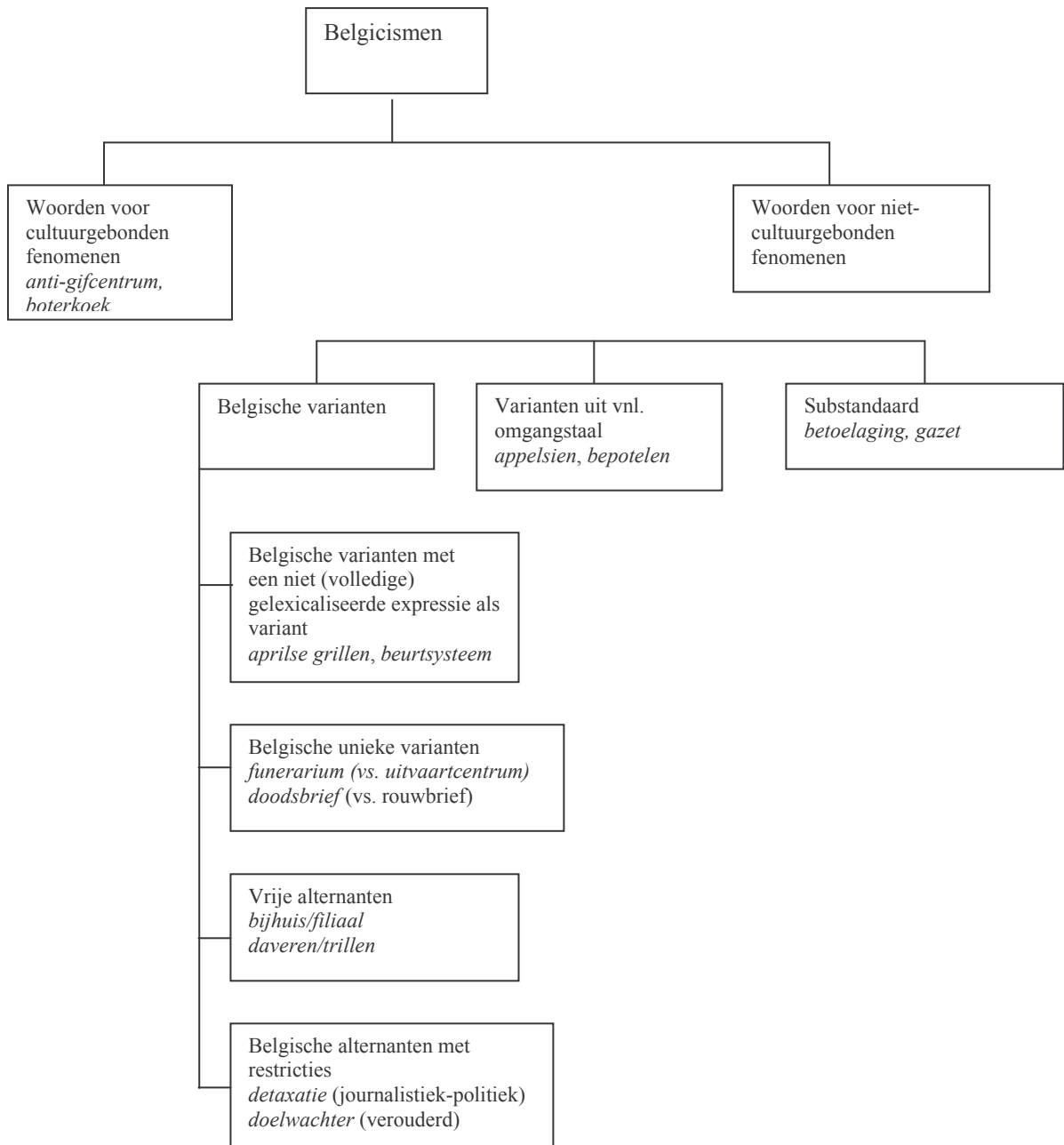
-Het bestand wil niet alleen een selectie van Belgicismen geven maar ook een methode om deze te identificeren en te valideren. De methode maakt gebruik van lexicografische bronnen, tekstcorpora en informanten. De methodologische gegevens vormen een onderdeel van het databestand.

NB. De definitieve documentatie bevat een toelichting met betrekking tot de manier waarop de data gebruikt en geïnterpreteerd zijn.

Doel

Een genuanceerde beschrijving geven van het typische Nederlandse woordgebruik in België door middel van de volgende 7 subcategorieën.

Subcategoriën: overzicht en toelichting



- **cultuurgebonden items**

woorden/uitdrukkingen die fenomenen benoemen die typisch zijn voor (de cultuur van) de samenleving waarin het Belgisch-Nederlands gebruikt wordt, nl. die van (Nederlandstalig) België. Soms hebben die termen een ‘tegenhanger’ in de cultuur van het Noorden (bv. ‘schepen’/’wethouder’). Voor elk van de items die tot deze categorie behoren, is er een verdere onderverdeling gemaakt naargelang van het domein waartoe ze kunnen behoren: politiek, juridische wereld, sociaal leven, sociale zekerheid enz.

- **vrije alternanten**

lexicale varianten die in het Nederlandstalige deel van België frequent voorkomen en meestal een Nederlands-Nederlands synoniem hebben. Beide termen zijn (ongeveer) gelijkwaardig en in België worden de twee mogelijkheden (min of meer) door elkaar gebruikt (bv. ‘dollekoeienziekte’/’gekekoeienziekte’).

- **omgangstalige varianten**

woorden die in Vlaanderen heel vaak worden gebruikt in de spreektaal, en ook wel worden geschreven, maar die door de Belgisch-Nederlandse spraakmakende gemeente niet als voorkeursvorm worden beschouwd, althans niet voor de schrijftaal. (bv. ‘(het gras) afrijden’). In de regel komen deze varianten niet in het Nederlands van Nederland voor.

- **substandaard**

Belgische varianten die wel voorkomen in het niet-dialectische taalgebruik van Nederlandstaligen (vooral in hun spreektaal), maar als niet correct worden beschouwd door prestigesprekers van het Nederlands in België (in die zin dat deze laatsten die varianten niet gebruiken en/of het gebruik ervan afkeuren) (bv. ‘camionette’ voor bestelwagen).

- **niet-gelexicaliseerde varianten**

items die niet variëren met een woord of uitdrukking in het Nederlands van Nederland, maar met een niet-gelexicaliseerde expressie. M.a.w. in het Nederlandse Nederlands kent men geen synoniem voor deze Belgische varianten, maar zal men ze omschrijven. Dit zijn de zogenaamde lexicale leemtes (in het NN) (bv. ‘vluchtsmisdrijf’).

- **alternanten met restricties**

items die alleen in het Belgische Nederlands voorkomen en die in een of andere zin (maar anders dan louter omgangstalig) gerestringeerd zijn (bv. ‘unief’ (informeel, vnl. studententaal) vs. universiteit).

- **unieke varianten**

woorden/uitdrukkingen die typisch zijn voor het Belgische Nederlands en in Nederland niet (of nauwelijks) voorkomen en die variëren met woorden/uitdrukkingen die typisch zijn voor het Nederlands in Nederland en die niet (of nauwelijks) voorkomen in het Belgische Nederlands (bv. ‘lintmeter’ vs. ‘centimeter’).

De subcategoratie is tot stand gekomen aan de hand van de informatie die over een bepaald item verzameld is (zie hieronder). Om het “type Belgicisme” nauwkeurig te kunnen bepalen zijn er een aantal algoritmes ontwikkeld. Bij iedere entry vindt men velden zoals “VERSCH(ueren) en “label-VS”, “GVD-label” en “(GVD-)restrictie”, NRC- en Roul.-freq(uentie), de bron (GVD, VdW, VUC of WCT), de cultuurgebondenheid en het domein de restricties op stilistisch/temporeel/sociaal en inhoudelijk vlak, de informant(en) en a- en b-scores en de variant (al dan niet gelexicaliseerd). Die velden zijn niet zomaar louter descriptief, maar maken deel uit van de procedure voor het bepalen van het “type Belgicisme”.

II. Toelichting bij de database

Formaat van de data: MS-ACCESS
Naam van de database : RBBN-1.0
Jaar van afronding: 2005
Naam van de tabel: belgicismen

Datacategorieën: veldnamen en -waarden

Hieronder volgt een opsomming van de velden van de database en een beschrijving van de inhoud en – indien relevant – een lijst van mogelijke waarden. Achter sommige veldnamen staat de code ‘M’ die als volgt geïnterpreteerd moet worden:

(M) methodologische data. Data die vooral van belang zijn om te komen tot een juiste subcategorisering van de items.

- **item**

entry: een woord of een uitdrukking

- **type item**

Hier zijn er twee mogelijke waarden: w(oord) of u(itdrukking). Woordgroepen krijgen de waarde ‘uitdrukking’.

- **sorteerlemma**

trefwoord waaronder het item dient te worden gerangschikt; alleen relevant als het om een uitdrukking gaat (type item = u)

- **woordklasse**

naamwoord	=	nw
werkwoord	=	ww
bijvoeglijk naamwoord	=	bnw
bijwoord	=	bw
bijvoeglijk naamwoord/bijwoord	=	bnw/bw
voorzetsel	=	vz
voegwoord	=	vw
telwoord	=	telw
voornaamwoord	=	vnw
tussenwerpsel	=	tw
werkwoordelijke groep	=	wwg
naamwoordelijke groep	=	nwg
bijvoeglijknaamwoordelijke groep	=	bnwg
voorzetselgroep	=	vzg
volzin	=	volzin
voegwoordelijke groep	=	vwg

nder het item dient te worden gerangschikt.

- **Type Belgicisme**

De subcategorieën volgens welke de items getypeerd worden:
Cultuurgebonden

vrije alternant
alternant met restricties
variant niet gelexicaliseerd
omgangstaal
substandaard
unieke variant
Zie hierboven voor een toelichting.

- **selectie BW**

Het RBBN is o.m. bedoeld als werkinstrument voor de samenstellers van bilinguale woordenboeken (cf. supra). De items (circa 1642) die zouden moeten opgenomen worden in een BW (= bilinguaal handwoordenboek met een omvang van 40.000 items) worden in dit veldje aangeduid met een vinkje. De defaultwaarde is 'neen'.

- **betekenisnummer**

Bij items met meer dan een Belgische betekenis, krijgen de verschillende betekenissen een volgnummer (bv. 'college': 1. vrije school voor middelbaar of hoger onderwijs aan jongens, 2. gebouw bij een universiteit waar studenten samenleven). De defaultwaarde is '1'.

- **definitie**

een korte betekenisomschrijving, doorgaans overgenomen uit een bestaande bron (die in het veld 'bron' vermeld wordt), als er in de betekenisomschrijving een asterisk voor een woord staat, betekent dit dat het woord in kwestie ook zelf als item in de database is opgenomen.

- **Bron**

Dit veld geeft de lexicografische bron aan waaraan het item is ontleend. Indien een woord afkomstig is uit een van de onderstaande bronnen dan is uit deze bron ook de definitie (zie veld *betekenis*) overgenomen.

GVD (Van Dale Groot Woordenboek der Nederlandse Taal)

VdW (Verboden op het werk te komen van Ton van der Wouden)

VUC (het V.U.-corpus)

WCT (Woordenboek voor Correct Taalgebruik) of eig. (eigen inbreng).

NB 1. Dit veld heeft een methodologische waarde (*aan welke bron is het item aan ontleend*) en een meta-lexicografische waarde (*wat is de bron van de definitie*)

2. In het veld *commentaar2004* staan de bronvermeldingen van items die uit andere dan een van de bovengenoemde bronnen voorkomen.

- **NRC-freq (M)**

De frequentie van het item volgens de Nederlandse component van het corpus INL-38miljoen (het zogenaamde Nederlandse ReferentieCorpus met een omvang van 22,5 miljoen woorden). Als het item 10 keer of meer voorkomt in het NRC, wordt het beschouwd als 'algemeen' en niet als belgicisme beschouwd.

- **Roul-freq (M)**

De frequentie van het item volgens het Roularta-corpus (het Belgische referentiecopus). Dit (Belgisch) corpus op CD-ROM bevat alle nummers van vijf tijdschriften (*Knack*, *Weekend-Knack*, *Cash!*, *Trends* en *Belgian Business and Industry*) uitgegeven door de Roularta Media Group van 1996 tot 1999. Bij de opbouw van het RBBN-bestand werd dit corpus gebruikt als Belgische tegenhanger van het Nederlands INL-corpus omdat de

Belgische component van het 38 miljoen INL-corpus te klein was qua omvang (2,6 miljoen woorden) en uitsluitend drie maanden van de krant *De Standaard* bevatte. Het ROULARTA-corpus telde in 1999 naar schatting 21,5 miljoen woorden. Anno 2005 telt het corpus 44.486.199 woorden.

De waarden die gehanteerd zijn, hebben een grenswaarde = 40 (item wordt beschouwd als belgicisme) en een min. waarde = 0. Wat betreft de variant van het item wordt een max. waarde = 78 gehanteerd en een min.waarde = 0. Dit is gedaan om praktische redenen: indien het gaat om vrije alternanten dan mag het niet zo zijn dat de variant van het item het dubbel aantal voorkomens van het item haalt. Als dat wel zo is, dan behoort het item tot een andere categorie ('alternant met restricties', 'omgangstaal' of 'substandaard').

- **google-nl/google-be (M)**

Voor een aantal items is een Google-frequentie genoteerd. Het betreft de items waarbij het type belgicisme in het originele bestand (1999) niet met zekerheid vastgelegd/vast te leggen was. De hoogste score is 40: op 40 sites is het aantal relevante hits geteld, het resultaat staat als frequentie genoteerd. Indien de Google-frequentie kleiner is dan 40, betekent dit dat het aantal websites kleiner was dan 40 en dan staat het reële aantal vermeld.

- **CGN-freq (M)**

De frequentie van ca. 100 spreektaalige items en uitdrukkingen volgens het CGN (= Corpus Gesproken Nederlands).

Nb. Helaas was er niet voldoende tijd de frequentiedata van alle items na te gaan.

- **GVD-label (M)**

= 'g' (gewestelijk) of 'b' (in België), alleen van toepassing bij entries die afkomstig zijn uit GVD.

- **restr(icties GVD) (M)**

Als GVD naast 'gew.' of 'in B.' nog andere labels toekent, staan die in bepaalde gevallen hier vermeld (doorgaans enkel wanneer dat invloed heeft op de subcategorisering, bv. als op basis van dat label het item in de categorie 'alternant met restricties' werd ondergebracht).

- **VERSCH (M)**

Komt het item voor in het Verschuieren groot encyclopedisch woordenboek (1996): 'ja' of 'nee'.

- **label-VS (M)**

'Ja' betekent hier dat het item door Verschuieren wordt gelabeld als 'in B.' (= in België) of 'Z.N.' (= Zuid-Nederlands) (of dat het in de definitie als zodanig wordt omschreven), 'Nee' betekent dat dat niet het geval is (maar dat sluit natuurlijk het voorkomen van andere labels niet uit).

- **cultuurgeb(onden) (M)**

De vraag naar de cultuurgebondenheid; 'ja' betekent dat het item een typisch Belgisch concept benoemt.

- **domein**

Volgende domeinen worden onderscheiden:

- politiek ('pol.'): namen voor Belgische politieke functies (bestendig afgevaardigde), instellingen (Vlaamse Raad) en gebruiken (alarmprocedure), alsook

- typisch Belgische politieke ideologieën (wallingantisme), etc.
- juridische wereld ('jur.'): rechtbank van eerste aanleg, vrederechter
 - onderwijs ('ond.'): ASO, blokperiode, inrichtende macht
 - financiële wereld ('fin.'): Bankcommissie, bevek
 - (socio-)economie ('sec.'): ABVV, délégué, Gesco
 - 'orde': de wereld van de ordehandhaving (cel, BOB), inclusief het leger (adjudant-chef)
 - administratie ('adm.'): namen voor overheidsinstellingen (ABOS, anti-gifcentrum) en functies (Conducteur van Bruggen en Wegen) allerhande, alsook officiële documenten (bewijs van goed zedelijk gedrag) e.d.
 - religieus leven ('rel.'): eremis, geuzengemeente
 - sociaal leven ('soc.'): heterogene categorie met namen voor allerlei gebruiken (doopsuiker, nieuwjaarsbrief, Bob), feestelijkheden (reuzenommegang), inclusief namen voor allerlei verenigingen (Chiro, KVLV), en concepten uit de sport (bevordering, flandrien)
 - culinaria ('cul.'): streekdranken (bolleke) en -'gerechten' (mattentaart, boterkoek, Antwerpse handjes)
 - historische termen ('his.'): blauwvoeterie, ambacht
 - ruimtelijke ordening ('r-o.'): landbouwzone, ambachtelijke zone
 - sociale zekerheid ('s-z.'): remgeld, OCMW
 - 'rest': een restcategorie met concepten die niet onder de voorafgaande domeinen ressorteren: dooibarelen, hollandiseren, P-trein

Een aantal items zouden in meer dan een categorie kunnen worden ondergebracht: 'Arbitragehof' is 'pol.' én 'jur.', 'rexisme' hoort zowel onder 'pol.' als onder 'his.', 'doopsuiker' is een gebruik, dus 'soc.', maar ook een 'gerecht', dus 'cul.', enz. Op het record staat echter telkens maar één domein aangegeven (bij twijfel is er dan doorgaans voor de meest 'ruime' categorie gekozen).

- **restricties (M)**

'Ja' betekent hier dat het item behoort tot de categorie 'alternanten met restricties'.

- **type-restricties (M)**

stilistisch ('stil'), temporeel ('temp'), sociaal ('soc') of inhoudelijk ('inh') waarde (restrictie)

- **waarde-restricties**

- stilistisch

formeel ('form.'): aflijvig, verbeiden

plechtig ('plechtig'): fiancé(e), wende

informeel ('inf.'): amigo, bib, micro

- temporeel

verouderd ('veroud.'): christen (adj.), refuge

modewoord ('modew.'): sans-papiers, hearing

- sociaal

journalistiek-politiek jargon ('j-p.'): culpabiliseren, penalisieren

studententaal ('stud.'): schacht

ambtelijke taal ('ambt.t.'): gemeentemandataris, hogestaand, wegenis

rechtstaal ('jur.'): mandaat

onderwijsjargon ('ond.'): reffectatie, valven

sporttaal ('sportt.'): brilscore

- inhoudelijk

‘sterker’: schandaliseren, onderhorige (heterogene groep van woorden die op een andere manier ‘sterker’, ‘pregnanter’ klinken dan hun varianten)

pejoratief (‘pej.’): tateren

schertsend (‘scherts.’): boefer, kaaskop, kiekenfretter

literaire taal/bijzondere stijl (‘lit.t.’): calvarietocht, tribulatie

in een culturele context (‘cult.’): Noord-Nederlander

- rest

een aantal items zijn gelabeld als ‘minder gebruikelijk’ (m.g.) om te behoeden dat ze in de categorie ‘omgangstaal/substandaard’ zouden terechtkomen, het gaat hier om items die gewoon veel minder gebruikelijk zijn dan hun varianten, zonder dat daarvoor een bepaalde reden (lees: restrictie) kan worden aangewezen (bv. olfactief vs. olfactorisch).

• **a-score trefwoord (M)**

Dit veld geeft de A-score van het item. De A-score is een score op een schaal van 0 tot 40 en is het antwoord van informanten (zie *informaten (bron)*) op de vraag: *Ken je dit woord (of uitdrukking)*.

Een item wordt als belgicisme geaccepteerd als het een A-score heeft die groter is dan of gelijk aan 20 en een B-score (zie *b-score trefwoord*) die groter is dan of gelijk aan 1.

• **b-score trefwoord (M)**

Dit veld geeft de B-score van het item. De B-score is een score op een schaal van 0 tot 2 en is het antwoord van informanten (zie *informanten (bron)*) op de vraag *Vind je dit woord (of deze uitdrukking) een goed AN-woord (of -uitdrukking)*.

Een item wordt als belgicisme geaccepteerd als het een A-score (zie *a-score trefwoord*) heeft die groter is dan of gelijk aan 20 en een B-score die groter is dan of gelijk aan 1.

• **informanten (bron) (M)**

Deygers (‘d’) of Coleman (‘c’):

Katrien Deygers heeft als licentiaatverhandeling aan de Universiteit Gent de labeling van Belgische woorden in de GVD (1992) bestudeerd. In het kader van die verhandeling heeft zij een uitgebreide enquête gehouden waarbij de informanten o.a. twee vragen moesten beantwoorden (zie A- en B-score).

T. Coleman heeft een uitbreiding gedaan van de Deygers-cijfers: hij heeft dezelfde vragen gesteld als in de enquête van K. Deygers maar voor andere woorden en uitdrukkingen.

De groep informanten bestond uit 20 Vlamingen gedifferentieerd, naar leeftijd, sexe, sociale klasse en regio van herkomst.

• **NN-variant (M)**

De NN variant(en) van het item. Als er geen waarde is ingevuld wil dit zeggen dat het item geen variant heeft in het NN.

• **Lexicalisatie van variant (M)**

Indicatie of de variant wel (ja) of niet (nee) gelexicaliseerd is.

• **Syntaxis (M)**

Bij een klein aantal items en hun varianten gaat het louter om een syntactisch verschil. Indien er sprake is van dit fenomeen, wordt dit in dit veld aangeduid met een vinkje. De defaultwaarde is hier ‘neen’. Bv. ‘in coma liggen vs. ‘in de coma liggen’; ‘bij klaarlichte dag vs. ‘op klaarlichte dag’; ‘een geluk met een ongeluk’ vs. ‘een geluk bij een ongeluk’ enz.

- **a-score variant (M)**

De a-score van de variant

Voor een toelichting bij a-score zie *a-score trefwoord*. In de definitieve documentatie volgt informatie over hoe deze score geïnterpreteerd dient te worden.

- **b-score variant (M)**

De b-score van de variant

Voor een toelichting bij b-score zie *b-score trefwoord*. In de definitieve documentatie volgt informatie over hoe deze score geïnterpreteerd dient te worden.

- **roul-freq van variant (M)**

De frequentie van de variant volgens het Roularta-corpus (het Belgische referentiecorpus). Zie ook het veld *roul-freq*.

- **opmerkingen**

- D: eventuele opmerkingen over de Deygers-cijfers

- R: hier worden additionele restricties gegeven en wordt ook het verband belicht tussen de verschillende typisch Belgische benamingen voor hetzelfde concept, bv. bij ‘cassatie’ vinden we ‘cassatie = informeel alternatief voor ‘Hof van Cassatie’.

- F: occasionele opmerkingen over frequenties: bv. als de Belgische variant vrij alterneert met zijn AN-tegenhanger, maar zelf toch veel vaker voorkomt (als bij ‘buitenwipper’ vs. ‘uitsmijter’), dan wordt dat hier aangegeven.

- O: de ‘o’ van opraaprocedure; als een item C- of S-woord (cultuurgebonden of substandaard) is, wordt dat hier vermeld: er zijn een aantal weinig frequente items die toch op basis van hun cultuurgebondenheid een plaats in het RBBN verdienen (de zogenaamde C-woorden; bv. ‘feest van de dynastie’), en een aantal items die weliswaar in de neutrale schrijftaal nauwelijks voorkomen, maar heel vaak worden gebruikt in de informele spreektaal (de zogenaamde S-woorden; bv. ‘elektrieker’). (Colleman 1999:33)

- **commentaar 2004**

Het RBBN is in twee periodes gemaakt de periode 1998/9 en de periode 2004/5 (zie inleiding) ; het merendeel van de meta-lexicografische data van de items die in de de tweede periode zijn toegevoegd aan het bestand, worden in dit veld vermeld. Gelijksortige informatie voor de woorden uit de eerste periode staan in de velden *bron, domein, restricties* enz. In de definitieve versie van het RBBN zullen – voor zover mogelijk - de metalexigrafische data van de woorden uit de eerste en de tweede periode geïntegreerd worden. LET OP: DE INFORMATIE DIE IN DIT VELD STAAT IS NIET OF NAUWELIJKS GESTRUCTUREERD; IN DE DEFINITIEVE VERSIE VAN HET DATABESTAND ZAL DAT WORDEN RECHTGEZET.

Het gaat om de volgende informatie:

Bron

Lijst N-DK (lijst Nederlands-Deens, een lijst gefilterd uit een groter geheel waarin enkel nog ‘vermeende’ belgicismen stonden)

Lijst VDL (lijst met woorden die in Van Dale (1999) als belgicisme zijn gelabeld)

DS1 en DS2 (twee verschillende bijlagen uit De Standaard met ‘vermeende’ belgicismen);

Ba (items uit het Vlaams-Nederlands Woordenboek van P. Bakema e.a.)

Lijst Moerdijk (lijst van ‘vermeende’ belgicismen opgesteld door A. Moerdijk e.a.)

TTV (items uit het archief van Taaltelefoon Vlaanderen, d.i. een dienst van het Ministerie van de Vlaamse Gemeenschap die beredeneerd en gemotiveerd taaladvies geeft over de Nederlandse taal)

BK (staat voor ‘Belgische kranten’; dit is een lijst met ‘vermeende’ belgicismen uit

(voornamelijk) Belgische kranten) en CT (items uit Correct Taalgebruik van Penninckx e.a.).

Elk van deze bronnen wordt in dit veld voorafgegaan door ‘Bron:’.

overige informatie: de overige informatie – indien het gaat om een nieuw item – wordt gescheiden van de bron door een liggend streepje. De verschillende soorten informatie worden binnen dit blok gescheiden door een puntkomma. Dit is ook het geval wanneer het niet gaat om een nieuw item. Het gaat hier om de volgende soorten informatie:

- **informatie m.b.t. de Google-frequentie:** soms kan de Google-frequentie voor verwarring zorgen. In dat geval wordt hier een verklaring gegeven (bv. ‘lastenboek’, ‘loontrekkende’); soms kan er helemaal geen Google-frequentie genoteerd worden (bv. het item heeft meerdere betekenissen en ook in combinatie met andere woorden kan er geen juiste hit gevonden worden: ‘leefbaar’, ‘leeggoed’); soms moet er ook in combinatie met andere woorden gezocht worden om een relevante hit te kunnen vinden. Deze woorden worden hier genoteerd (bv. ‘klasseren’); ten slotte wordt het item soms helemaal niet gevonden in Google.
- **extra label:** af en toe worden in dit veld ook extra labels genoteerd, zoals bijvoorbeeld formeel (‘lasterlijke aantijging’), informeel (‘avondkleed’), spreektaal (‘afzakken (naar)’), ... Ze duiden allemaal op het (eventuele) spreektaalige aspect van het betreffende item. Een item kan in vele gevallen twee labels, twee types belgicismen toegekend krijgen. Toch kan slechts één label genoteerd worden in het veldje “type belgicisme”. De informatie die dan verloren gaat, wordt eveneens in het veld “commentaar 2004” genoteerd (bv. ‘beschermde werkplaats’, ‘de weerstand’).
- **Vakterminologie** sporttaal (‘thuismatch’), culinaire terminologie (‘loze vinken’), politiek jargon (‘voogdijminister’), vergaderterminologie (‘verontschuldigd zijn’), socio-culturele sector (‘wijkmeester’), onderwijsterminologie (‘thesisjaar’), financiële sector (‘verwijlinterest’), juridische terminologie (‘onverdeeldheid’), studententaal (‘brosser’), literair taalgebruik (‘traagzaam’), administratieve sector (‘onroerende voorheffing’) ...
- **gebruik van het item (frequentie, voorbeelden, contexten ...):** in dit blokje staat informatie over het gebruik en de frequentie van het betreffende item: in sommige gevallen wordt het item minder gebruikt dan een aantal jaren geleden (bv. ‘halftijds’), is het item verouderd (bv. ‘maturiteitsdiploma’) of in onbruik geraakt (bv. ‘ministerin’) of is het sinds zeer kort van naam veranderd (bv. ‘RUCA’, ‘UFSIA’ en ‘UIA’ hebben nu allemaal de naam UA, Universiteit Antwerpen). Daarnaast wordt er ook informatie gegeven over het feit dat het item meer gebruikt wordt in België dan in Nederland (bv. ‘kernkabinet’) of enkel gebruikt wordt in België (bv. ‘het laken naar zich toe trekken’) en in welke contexten. Af en toe vind je ook de afkorting ‘m.g.’ (= minder gebruikelijk). Dit wil zeggen dat het item bijvoorbeeld minder gebruikt wordt dan zijn variant (bv. ‘repertoriëren’). Soms staat er ook ‘X’ is frequenter dan ‘Y’ (waarbij X = item en Y = variant) (bv. ‘steenwol’). Ten slotte vind

je hier eveneens voorbeelden van contexten (gevonden op Google) (bv. ‘iet of wat’) waarin het item zeer vaak gebruikt wordt of worden er voorbeeldzinnnetjes gemaakt, bijvoorbeeld wanneer het item voorkomt in een veel gebruikte combinatie van woorden (bv. ‘grieventrommel’, ‘goesting’, ‘ophebben’).

- **commentaar 2000**

commentaar zoals hij is ingegeven door Coleman in het originele bestand (1999). Dit veld bevat informatie over o.a. de cultuurgebondenheid die genuanceerd dient te worden, problemen bij het kiezen tussen verschillende varianten enz. De betreffende commentaar wordt gevolgd door ‘(2000)’.

verwijzingen: indien in het RBBN-bestand eveneens de formelere resp. informelere, de verouderde of meer officiële variant van een item voorkomt, dan wordt hiernaar verwezen door de betreffende variant in dit veld te noteren, (eventueel) voorafgegaan door ‘zie’ (bv. ‘intensieve zorgen’, ‘wegcode’, ‘zittijd’).

extra: de overige informatie komt slechts enkele malen voor en probeer ik hier samen te vatten: soms is de relatie tussen item en variant niet altijd even duidelijk. In die gevallen wordt in dit veld een Google-frequentie genoteerd van een of meerdere varianten. Zo wordt er een duidelijker beeld gecreëerd wanneer men twijfelt tussen bijvoorbeeld een ‘vrije alternant’ en een ‘unieke variant’ (bv. ‘afbraakwerken’); verder alterneren varianten niet altijd volledig met het betreffende item. In dat geval wordt er eveneens een verklaring gegeven in het “commentaar 2004”-veld (bv. ‘kort op de bal spelen’); daarnaast zijn er items die op verschillende manieren geschreven kunnen worden. De spellingsvarianten van deze items vind je hier terug (bv. ‘délégué’ en ‘delegee’); ten slotte zijn er een aantal items die enkel of vaker in hun meervoudsvorm worden gebruikt (bv. ‘naakt ontslag’), ook dat wordt hier genoteerd.